1 Простейшая задача вариационного исчисления

Сначала напомним, что

1. $C([t_0,\, t_1],\, \mathbb{R}^n)$ — это пространство непрерывных функций $x : [t_0,\, t_1] \to \mathbb{R}^n$; на нем вводится норма
$$\|x\|_C = \max_{t\in[t_0,\, t_1]} \max_{1\le j\le n} |x_j(t)|$$
(здесь $x(t) = (x_1(t),\, \dots,\, x_n(t))$);

2. $C^1([t_0,\, t_1],\, \mathbb{R}^n)$ — это пространство непрерывных функций $x : [t_0,\, t_1] \to \mathbb{R}^n$; на нем вводится норма $\|x\|_{C^1} = \max\{\|x\|_C,\, \|\dot{x}\|_C\}$.

Пусть $L : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ — непрерывное отображение. Тогда определено отображение

$$\mathcal{L} : C^1([t_0,\, t_1],\, \mathbb{R}^n) \to \mathbb{R}, \quad \mathcal{L}(x) = \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt$$

($x(t)$ и $\dot{x}(t)$ — это элементы $\mathbb{R}^n$ для любого $t \in [t_0,\, t_1]$; функция $t \mapsto L(t,\, x(t),\, \dot{x}(t))$ непрерывна по теореме о непрерывности сложной функции).

Рассмотрим следующую задачу на минимум:

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt \to \inf, \quad x(t_0) = x_{(0)}, \quad x(t_1) = x_{(1)}. \tag{1}$$

Функция $x \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$ называется допустимой, если $x(t_0) = x_{(0)}$ и $x(t_1) = x_{(1)}$.

Функция $\hat{x} \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$ называется точкой локального минимума для задачи (1), если она допустимая и существует $\varepsilon > 0$ такое, что для любой допустимой функции x такой, что $\|x-\hat{x}\|_{C^1} < \varepsilon$ выполнено $\mathcal{L}(x) \ge \mathcal{L}(\hat{x})$.

Необходимое условие локального минимума формулируется при дополнительном условии на гладкость функции L. А именно, предполагаем, что функция $(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1,\, \dots,\, \eta_n) \mapsto L(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1,\, \dots,\, \eta_n)$

имеет частные производные по каждой переменной ξ_i и η_i $(1 \le i \le n)$, при этом отображения

$$(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_n) \mapsto L_{\xi_i}(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_n)$$

и

$$(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_n) \mapsto L_{\eta_i}(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_n)$$

непрерывны.

Далее будем обозначать: $L_{\xi_i} =: L_{x_i}$, $L_{\eta_i} =: L_{\dot{x}_i}$, $L_x = (L_{x_1},\, \ldots,\, L_{x_n})$, $L_{\dot{x}} = (L_{\dot{x}_1},\, \ldots,\, L_{\dot{x}_n})$. Для n-мерных векторов v, w через vw будем обозначать их скалярное произведение в $\mathbb{R}^n$.

Обозначим

$$C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n) = \{h \in C^1([t_0,\, t_1], \mathbb{R}^n) :\ h(t_0) = h(t_1) = 0\}.$$

Предложение 1. *Пусть $\hat{x}$ — точка локального минимума в задаче* (1). *Тогда для любой функции $h \in C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n)$ выполнено*

$$\int_{t_0}^{t_1} \Big(L_x(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) h(t) + L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) \dot{h}(t) \Big)\, dt = 0. \qquad (2)$$

Доказательство. Рассмотрим функцию $\varphi(\lambda) = \mathcal{L}(\hat{x} + \lambda h)$. Вычислим ее производную в точке $\lambda = 0$:

$$\varphi'(0) = \lim_{\lambda \to 0} \int_{t_0}^{t_1} \frac{L(t,\, \hat{x}(t) + \lambda h(t),\, \dot{\hat{x}}(t) + \lambda \dot{h}(t)) - L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))}{\lambda}\, dt.$$

По теореме о производной композиции, для любого $t \in [t_0,\, t_1]$

$$\lim_{\lambda \to 0} \frac{L(t,\, \hat{x}(t) + \lambda h(t),\, \dot{\hat{x}}(t) + \lambda \dot{h}(t)) - L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))}{\lambda} =$$

$$= L_x(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) h(t) + L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) \dot{h}(t).$$

Остается переставить предел и интеграл. Это можно сделать по теореме Лебега о мажорируемой сходимости. В самом деле, пусть $\psi(\lambda,\, t) =$

$L(t,\, \hat{x}(t)+\lambda h(t),\, \dot{\hat{x}}(t)+\lambda \dot{h}(t))$. Из теоремы Лагранжа следует, что для любого $t \in [t_0,\, t_1]$ существует число $\theta(t) \in [0,\, 1]$ такое, что $\psi(\lambda,\, t)-\psi(0,\, t) = \psi'_\lambda(\lambda\theta(t),\, t)\lambda$, откуда

$$\frac{L(t,\, \hat{x}(t)+\lambda h(t),\, \dot{\hat{x}}(t)+\lambda \dot{h}(t)) - L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))}{\lambda} =$$

$$= L_x(t,\, \hat{x}(t)+\lambda\theta(t)h(t),\, \dot{\hat{x}}(t)+\lambda\theta(t)\dot{h}(t))h(t)+$$
$$+L_{\dot{x}}(t,\, \hat{x}(t)+\lambda\theta(t)h(t),\, \dot{\hat{x}}(t)+\lambda\theta(t)\dot{h}(t))\dot{h}(t) =: f_\lambda(t).$$

Так как функции x, $\dot{x}$, h, $\dot{h}$ ограничены, а функции L_x и $L_{\dot{x}}$ непрерывны, то существует $M > 0$ такое, что $|f_\lambda(t)| \leq M$ для любого $t \in [t_0,\, t_1]$, $\lambda \in [-1,\, 1]$. Значит, $g(t) \equiv M$ будет интегрируемой мажорантой для f_λ. $\square$

Оказывается, (2) эквивалентно некоторой системе обыкновенных дифференциальных уравнений, называющейся уравнениями Эйлера – Лагранжа.

Сначала напомним одно из эквивалентных определений абсолютно непрерывной функции.

Определение 1. *Функция $f : [t_0,\, t_1] \to \mathbb{R}$ называется абсолютно непрерывной, если существует функция $g \in L_1[t_0,\, t_1]$ такая, что $f(t) = f(t_0) + \int\limits_{t_0}^{t} g(s)\, ds$.*

Абсолютно непрерывная функция почти всюду дифференцируема и $f'(t) = g(t)$ п.в.

Пространство абсолютно непрерывных функций на $[t_0,\, t_1]$ обозначим через $AC[t_0,\, t_1]$.

Пусть $f,\, \varphi \in AC[t_0,\, t_1]$. Тогда выполнена формула интегрирования по частям:

$$\int\limits_{t_0}^{t_1} f'(t)\varphi(t)\, dt = -\int\limits_{t_0}^{t_1} f(t)\varphi'(t)\, dt + f(t)\varphi(t)\big|_{t_0}^{t_1}.$$

Предложение 2. *Равенство* (2) *для любой функции* $h \in C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n)$ *выполнено тогда и только тогда, когда функция* $t \mapsto L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ *абсолютно непрерывна и выполнены уравнения Эйлера–Лагранжа*

$$-\frac{d}{dt}L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) + L_{x_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) = 0, \quad 1 \le j \le n; \tag{3}$$

в векторной форме эта система записывается как

$$-\frac{d}{dt}L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) + L_{x}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) = 0.$$

Доказательство $\Leftarrow$**.** Пусть $L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ абсолютно непрерывны и выполнено (3). Тогда для любого $h \in C^1_{0,0}([t_0,\, t_1],\, \mathbb{R}^n)$

$$\int\limits_{t_0}^{t_1} \sum_{j=1}^{n} \left(-\frac{d}{dt}L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) + L_{x_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) \right) h_j(t)\, dt = 0.$$

Проинтегрировав по частям $\int\limits_{t_0}^{t_1} \frac{d}{dt}L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) h_j(t)\, dt$ и учитывая условия $h_j(t_0) = h_j(t_1) = 0$, получаем (2).

Тем же способом доказать утверждение в другую сторону не получится, так как мы заранее не знаем, можно ли интегрировать по частям. Нам понадобится **лемма Дюбуа-Реймона**. Сразу докажем ее обобщенную форму.

Для нормированного пространства X через X^* обозначаем сопряженное пространство (т.е. пространство линейных непрерывных функционалов на X).

Обобщенная лемма Дюбуа-Реймона. Пусть $f \in (C[t_0,\, t_1])^*$, при этом $f(\dot{h}) = 0$ для любой функции $h \in C^1_{0,0}[t_0,\, t_1]$. Тогда существует число $c \in \mathbb{R}$ такое, что

$$f(x) = c \int\limits_{t_0}^{t_1} x(t)\, dt, \quad x \in C[t_0,\, t_1].$$

Доказательство. Идея такая же, как у похожей теоремы из функционального анализа: если обобщенная производная равна 0, то обобщенная функция — константа.

Пусть $x \in C[t_0, t_1]$, $\int\limits_{t_0}^{t_1} x(t)\,dt = 0$. Положим $h(t) = \int\limits_{t_0}^{t} x(s)\,ds$. Тогда $h \in C^1[t_0, t_1]$, $h(t_0) = h(t_1) = 0$. По условию, $f(x) = f(\dot h) = 0$.

Пусть теперь $x \in C[t_0, t_1]$ — произвольная функция. Пусть $\xi \in C[t_0, t_1]$, $\int\limits_{t_0}^{t_1} \xi(t)\,dt = 1$. Положим

$$y(t) = x(t) - \xi(t)\int\limits_{t_0}^{t_1} x(s)\,ds.$$

Тогда $\int\limits_{t_0}^{t_1} y(t)\,dt = 0$. В силу доказанного, $f(y) = 0$. Значит, $f(x) = f(\xi)\int\limits_{t_0}^{t_1} x(s)\,ds$. Осталось положить $c = f(\xi)$. $\diamond$

Следствие. Если функционал f удовлетворяет условиям леммы Дюбуа-Реймона и имеет вид

$$f(x) = \int\limits_{t_0}^{t_1} y(t)x(t)\,dt,$$

где $y \in L_1[t_0, t_1]$, то $y(t) = c$ п.в.

Доказательство. Достаточно доказать, что $\int\limits_{t_0}^{t} y(s)\,ds = c(t - t_0)$ для любого $t \in [t_0, t_1]$. Это получается предельным переходом в равенстве

$$\int\limits_{t_0}^{t_1} y(s)x_n(s)\,ds = \int\limits_{t_0}^{t_1} cx_n(s)\,ds,$$

где $x_n \in C[t_0, t_1]$, $x_n \to \chi_{[t_0, t]}$ п.в., $x_n(s) \in [0, 1]$ для любого $s \in [t_0, t_1]$. $\diamond$

Доказательство $\Rightarrow$. Пусть выполнено (2) для любой функции $h \in C^1_{0,0}([t_0, t_1], \mathbb{R}^n)$. Положим $B_j(t) = \int\limits_{t_0}^{t} L_{x_j}(s, \hat x(s), \dot{\hat x}(s))\,ds$, $B = (B_1, \ldots, B_n)$. Проинтегрируем по частям в (2) $L_{x_j}(t, \hat x(t), \dot{\hat x}(t))h_j(t)$ для всех $j = 1, \ldots, n$, учитывая, что $h(t_0) = h(t_1) = 0$, и получим:

$$\int\limits_{t_0}^{t_1} (L_{\dot x}(t, \hat x(t), \dot{\hat x}(t)) - B(t))\dot h(t)\,dt = 0. \qquad (4)$$

Зафиксируем $j \in \overline{1,\, n}$ и рассмотрим h такие, что $h_i = 0$ при $i \neq j$. Из (4) и следствия из леммы Дюбуа-Реймона следует, что

$$L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) - B_j(t) \equiv \text{const}, \quad 1 \le j \le n.$$

Значит, $L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ абсолютно непрерывны и выполнено (3).

Итак, доказана

Теорема 1. *Пусть $\hat{x}$ — точка локального минимума в (1). Тогда выполнены уравнения Эйлера – Лагранжа (3).*

Допустимая функция $\hat{x}$, удовлетворяющая (3), называется допустимой экстремалью.

Замечание. Мы рассмотрели классическую постановку задачи на пространстве C^1. Можно рассмотреть постановку этой задачи на пространстве липшицевых функций $\mathrm{Lip}([t_0,\, t_1],\, \mathbb{R}^n)$. Напомним, что каждая липшицева функция абсолютно непрерывна и ее производная (которая существует п.в.) существенно ограничена. В качестве нормы на $\mathrm{Lip}([t_0,\, t_1],\, \mathbb{R}^n)$ берется

$$\|x\|_{\mathrm{Lip}} = \max\{\|x\|_C,\, \|\dot{x}\|_{L_\infty}\}.$$

Если L непрерывна, то функция $t \mapsto L(t,\, x(t),\, \dot{x}(t))$ принадлежит $L_\infty[t_0,\, t_1]$, так что функционал $\mathcal{L}$ корректно определен. Если при этом L_x и $L_{\dot{x}}$ существуют и непрерывны, то все утверждения выше верны (в предложении 1 снова можно применить теорему Лебега, в предложении 2 можно применить следствие из леммы Дюбуа-Реймона и проинтегрировать по частям; значит, теорема 1 тоже верна).

Пример: гармонический осциллятор. Рассмотрим задачу

$$\mathcal{L}(x) := \int\limits_0^{T_0} (\dot{x}^2 - c^2x^2)\, dt \to \inf, \quad x(0) = x(T_0) = 0;$$

здесь $c > 0$, $T_0 > 0$ — фиксированные параметры. Тогда $L_{\dot{x}} = 2\dot{x}$, $L_x = -2c^2x$; уравнение Эйлера имеет вид $-\frac{d}{dt}(2\dot{x}) - 2c^2x = 0$, т.е. $\ddot{x} + c^2x = 0$. Решение этого уравнения, удовлетворяющее условию $x(0) = 0$, имеет вид $x(t) = a\sin ct$. Если $cT_0 \neq \pi k$ ($k \in \mathbb{N}$), то условию $x(T_0) = 0$ удовлетворяет только $x = 0$; если $cT_0 = \pi k$, $k \in \mathbb{N}$, то $x(t) = a\sin ct$ — допустимая экстремаль при всех $a \in \mathbb{R}$.

Выясним, является ли $\hat{x} = 0$ точкой локального или глобального минимума (исследование решения в общем случае будет на семинарах). Для этого используем следующий прием.

Пусть $\omega \in C^1[0,\, T_0]$. Тогда $\int\limits_0^{T_0} \frac{d}{dt}(\omega x^2)\, dt = \omega x^2|_0^{T_0} = 0$, если $x \in C^1_{0,0}[0,\, T_0]$. Значит,

$$\int\limits_0^{T_0} (\dot{x}^2 - c^2x^2)\, dt = \int\limits_0^{T_0} (\dot{x}^2 - c^2x^2 - \dot{\omega}x^2 - 2\omega x\dot{x})\, dt.$$

Подберем ω так, чтобы $\dot{x}^2 - c^2x^2 - \dot{\omega}x^2 - 2\omega x\dot{x}$ было полным квадратом: $\dot{x}^2 - c^2x^2 - \dot{\omega}x^2 - 2\omega x\dot{x} = (\dot{x} - \omega x)^2$, т.е. $-c^2 - \dot{\omega} = \omega^2$. (Тогда $\int\limits_0^{T_0} (\dot{x}^2 - c^2x^2)\, dt = \int\limits_0^{T_0} (\dot{x} - \omega x)^2\, dt \geq 0$.) Из дифференциального уравнения получаем, что $\omega = c \operatorname{ctg} c(t - t_*)$.

1. Если $cT_0 < \pi$, то можно выбрать t_* так, чтобы функция $c \operatorname{ctg} c(\cdot - t_*)$ была гладкой на $[0,\, T_0]$; значит, рассуждения выше можно корректно провести и 0 — точка глобального минимума.

2. Пусть $cT_0 > \pi$. Возьмем $x(t) = a \sin \frac{\pi t}{T_0}$. Тогда x — допустимая функция, но $\mathcal{L}(x) < 0$.

3. Пусть $cT_0 = \pi$. Покажем, что $\int\limits_0^{\pi/c} (\dot{x}^2 - c^2x^2)\, dt = \int\limits_0^{\pi/c} (\dot{x} - x \cdot c \operatorname{ctg} ct)^2\, dt \geq 0$. Второе соотношение очевидно, а первое проверяем с помощью интегрирования по частям:

$$\int\limits_0^{\pi/c} (\dot{x} - x \cdot c \operatorname{ctg} ct)^2\, dt = \int\limits_0^{\pi/c} (\dot{x}^2 - 2\dot{x}x \cdot c \operatorname{ctg} ct + c^2x^2 \operatorname{ctg}^2 ct)\, dt =$$

$$= \int\limits_0^{\pi/c} (\dot{x}^2 + c^2x^2 \operatorname{ctg}^2 ct)\, dt - x^2 \cdot c \operatorname{ctg} ct|_0^{\pi/c} + \int\limits_0^{\pi/c} (x^2 \cdot c^2(-1 - \operatorname{ctg}^2 ct))\, dt =$$

$$= \int\limits_0^{\pi/c} (\dot{x}^2 - c^2x^2)\, dt - x^2 \cdot c \operatorname{ctg} ct|_0^{\pi/c}.$$

Остается показать, что если $x \in C^1_{0,0}[0,\, \pi/c]$, то $\lim\limits_{t\to 0+} x^2(t)\operatorname{ctg} ct = \lim\limits_{t\to\pi/c-} x^2(t)\operatorname{ctg} ct = 0$. Вычислим первый предел (второй аналогичный). Так как $x(0) = 0$ и x дифференцируема в нуле, то $x(t) = O(t)$, $t \to 0+$. Значит, $x^2(t)\operatorname{ctg} ct = O(t^2)O(t^{-1}) \underset{t\to 0}{\to} 0+$.

Вывод. Если $cT_0 \le \pi$, то 0 — точка глобального минимума. Если $cT_0 > \pi$, то 0 не является точкой локального минимума.

Отметим также, что если взять $x(t) = a\sin\frac{\pi k t}{T_0}$ при достаточно большом $k \in \mathbb{N}$, то $\mathcal{L}(x) > 0$. Поэтому локального максимума нет ни при каких значениях параметров.

Пример: геодезические на римановом многообразии. Пусть M — гладкое многообразие, g_{ij} — риманова метрика на M, γ — гладкая кривая на M. Если образ γ лежит в одной карте и в этой карте параметризация имеет вид $\gamma : t \mapsto x(t)$, то функционал длины кривой имеет вид

$$l(\gamma) = \int\limits_{t_0}^{t_1} \sqrt{g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t)}\, dt$$

(по повторяющимся индексам проводится суммирование). Этот интеграл не зависит от выбора параметризации кривой и локальных координат.

Докажем, что экстремалями для функционала длины кривой являются геодезические. Пусть $\hat{x}$ — экстремаль. Можем считать, что параметризация этой кривой пропорциональна натуральной, т.е. $g_{ij}(\hat{x}(t))\dot{\hat{x}}^i(t)\dot{\hat{x}}^j(t) = c$. Имеем

$$L_{\dot{x}^k} = \frac{g_{kj}(\hat{x})\dot{\hat{x}}^j}{\sqrt{g_{ij}(\hat{x})\dot{\hat{x}}^i\dot{\hat{x}}^j}} = \frac{1}{\sqrt{c}} g_{kj}(\hat{x})\dot{\hat{x}}^j,$$

$$L_{x^k} = \frac{\frac{\partial}{\partial x^k} g_{ij}(\hat{x})\dot{\hat{x}}^i\dot{\hat{x}}^j}{2\sqrt{g_{ij}(\hat{x})\dot{\hat{x}}^i\dot{\hat{x}}^j}} = \frac{1}{2\sqrt{c}}\frac{\partial}{\partial x^k} g_{ij}(\hat{x})\dot{\hat{x}}^i\dot{\hat{x}}^j.$$

Значит, уравнения Эйлера–Лагранжа имеют вид

$$-\frac{d}{dt}(g_{kj}(\hat{x})\dot{\hat{x}}^j) + \frac{1}{2}\frac{\partial}{\partial x^k} g_{ij}(\hat{x})\dot{\hat{x}}^i\dot{\hat{x}}^j = 0.$$

Если при этом $\hat{x} \in C^2$, то

$$-g_{kj}(\hat{x})\ddot{\hat{x}}^j - \frac{\partial}{\partial x_i} g_{kj}(\hat{x})\dot{\hat{x}}^i\dot{\hat{x}}^j + \frac{1}{2}\frac{\partial}{\partial x^k} g_{ij}(\hat{x})\dot{\hat{x}}^i\dot{\hat{x}}^j = 0.$$

Поменяв во втором слагаемом индексы i и j, получим эквивалентное уравнение (так как по i, j проводится суммирование):

$$-g_{kj}(\hat{x})\ddot{\hat{x}}^j - \frac{\partial}{\partial x_j} g_{ki}(\hat{x})\dot{\hat{x}}^j\dot{\hat{x}}^i + \frac{1}{2}\frac{\partial}{\partial x^k} g_{ij}(\hat{x})\dot{\hat{x}}^i\dot{\hat{x}}^j = 0.$$

Сложим эти два уравнения и поделим на 2:

$$-g_{kj}(\hat{x})\ddot{\hat{x}}^j - \frac{1}{2}\left(\frac{\partial}{\partial x_i} g_{kj}(\hat{x}) + \frac{\partial}{\partial x_j} g_{ki}(\hat{x}) - \frac{\partial}{\partial x^k} g_{ij}(\hat{x})\right)\dot{\hat{x}}^i\dot{\hat{x}}^j = 0.$$

Умножим на матрицу $g^{km}(\hat{x})$, обратную к $g_{kj}(\hat{x})$, и на -1:

$$\ddot{\hat{x}}^m + \frac{1}{2} g^{km}(\hat{x})\left(\frac{\partial}{\partial x_i} g_{kj}(\hat{x}) + \frac{\partial}{\partial x_j} g_{ki}(\hat{x}) - \frac{\partial}{\partial x^k} g_{ij}(\hat{x})\right)\dot{\hat{x}}^i\dot{\hat{x}}^j = 0;$$

вспомнив, как выражаются символы Кристоффеля через риманову метрику, получаем $\ddot{\hat{x}}^m + \Gamma^m_{ij}\dot{\hat{x}}^i\dot{\hat{x}}^j = 0$. Это и есть уравнение геодезической.

Когда можно заранее понять, что допустимая экстремаль будет C^2-гладкой?

Предложение 3. *Пусть $L \in C^2([t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n)$, при этом матрица $L_{\dot{x}\dot{x}}(t, \xi, \eta)$ невырожденная при всех $(t, \xi, \eta) \in [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n$. Пусть $\hat{x} \in C^1([t_0, t_1], \mathbb{R}^n)$ — решение уравнения Эйлера. Тогда $\hat{x} \in C^2([t_0, t_1], \mathbb{R}^n)$.*

Доказательство. Так как матрица $L_{\dot{x}\dot{x}}$ невырожденная и $L_{\dot{x}} \in C^1$, то уравнение $L_{\dot{x}}(t, \xi, \eta) = p$ локально разрешимо относительно η по теореме о неявной функции, при этом решение $\eta(t, \xi, p)$ будет C^1-гладким.

Функция $t \mapsto L_x(t, \hat{x}(t), \dot{\hat{x}}(t))$ непрерывна, поэтому ее первообразная $\varphi \in C^1[t_0, t_1]$; в силу уравнения Эйлера, $L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) = \varphi(t)$. Отсюда $\dot{\hat{x}}(t) = \eta(t, \hat{x}(t), \varphi(t))$. Правая часть C^1-гладкая. Значит, $\dot{\hat{x}} \in C^1([t_0, t_1], \mathbb{R}^n)$, откуда $\hat{x} \in C^2([t_0, t_1], \mathbb{R}^n)$. □

Замечание. Мы доказали, что экстремаль для функционала длины кривой является геодезической, в предположении, что она C^2-гладкая; при этом параметризация выбиралась пропорциональной натуральной. Для самого функционала длины матрица $L_{\dot{x}\dot{x}}$ не будет невырожденной. Но можно заметить, что при пропорциональной натуральной параметризации уравнения Эйлера–Лагранжа для функционала длины будут такими же, как для функционала

$$\mathcal{L}(x) = \int\limits_{t_0}^{t_1} g_{ij}(x)\dot{x}^i\dot{x}^j\,dt.$$

А для $L = g_{ij}(x)\dot{x}^i\dot{x}^j$ имеем $L_{\dot{x}\dot{x}} = (g_{ij}(x))$. Это положительно определенная и, в частности, невырожденная матрица. Значит, экстремали для функционала длины действительно C^2-гладкие.

Некоторые контрпримеры. Пространство $C^1([t_0,\,t_1],\,\mathbb{R}^n)$ не всегда является естественным для простейшей задачи вариационного исчисления. Это показывают следующие два примера.

Задача 1. Доказать, что в задаче

$$\int\limits_0^1 t^{1/2}\dot{x}^2\,dt \to \inf, \quad x(0) = 0, \quad x(1) = 1$$

точки локального минимума в пространстве $C^1[0,\,1]$ не существует; найти точку глобального минимума для той же задачи в пространстве

$$W = \left\{ f \in AC[0,\,1] : \int\limits_0^1 t^{1/2}\dot{x}^2\,dt < \infty \right\}.$$

Пусть $x : [t_0,\,t_1] \to \mathbb{R}$. Скажем, что x липшицева, если существует $C > 0$ такое, что для любых $t,\,s \in [t_0,\,t_1]$ выполнено $|x(t) - x(s)| \le C|t - s|$. Липшицева функция абсолютно непрерывна и поэтому п.в. дифференцируема.

Задача 2. 1) Доказать, что в задаче

$$\int\limits_0^1 (1 - \dot{x}^2)^2\,dt \to \inf, \quad x(0) = 0, \quad x(1) = 0$$

точки локального минимума в пространстве $C^1[0,\,1]$ не существует, при этом точная нижняя грань функционала равна 0; в пространстве липшицевых функций минимум достигается.

2) Доказать, что в задаче

$$\int\limits_0^1 ((1-\dot{x}^2)^2 + x^2)\,dt \to \inf, \quad x(0)=0, \quad x(1)=0$$

точки глобального минимума в пространстве липшицевых функций не существует, при этом точная нижняя грань функционала равна 0.

Задача 3. Рассмотрим функционал $J : AC[0,\,1] \to \mathbb{R} \cup \{+\infty\}$,

$$J(x) = \int\limits_0^1 \dot{x}^6 \cdot (x - t^{1/3})^2\,dt.$$

Показать, что $\inf\{J(x) :\ x \in AC[0,\,1],\ x(0)=0,\ x(1)=1\} < \inf\{J(x) :\ x \in \mathrm{Lip}[0,\,1],\ x(0)=0,\ x(1)=1\}$.

Уравнения Штурма–Лиувилля с сингулярным потенциалом: формализация через уравнения Эйлера–Лагранжа.

На дифференциальных уравнениях изучалась задача:

$$-\ddot{x} + v(t)x = \lambda x, \quad x(t_0) = x(t_1) = 0.$$

Нетрудно заметить, что это уравнение Эйлера–Лагранжа для

$$\mathcal{L}(x) = \int\limits_{t_0}^{t_1} (\dot{x}^2 + v(t)x^2 - \lambda x^2)\,dt.$$

В физике (например, в квантовой механике) возникают примеры, где $v(t)$ — обобщенная функция (например, содержит дельта-функцию в качестве слагаемого). В 1999 г. А.А. Шкаликов и А.М. Савчук дали несколько эквивалентных способов формализации таких уравнений для случая, когда $v = V'$ (обобщенная производная), $V \in L_2[t_0,\,t_1]$. Один из них можно описать через уравнения Эйлера–Лагранжа.

Сначала преобразуем функционал $\mathcal{L}$ в случае $v \in L_1[t_0,\,t_1]$ (сделаем интегрирование по частям):

$$\int\limits_{t_0}^{t_1} (\dot{x}^2 + v(t)x^2 - \lambda x^2)\,dt = \int\limits_{t_0}^{t_1} (\dot{x}^2 - 2V(t)x\dot{x} - \lambda x^2)\,dt.$$

А теперь рассмотрим функционал

$$\mathcal{L}(x) = \int\limits_{t_0}^{t_1} (\dot{x}^2 - 2V(t)x\dot{x} - \lambda x^2)\, dt$$

для $V \in L_2[t_0, t_1]$. Для него естественная область определения — пространство Соболева

$$W_2^1[t_0, t_1] = \{x \in AC[t_0, t_1] : \dot{x} \in L_2[t_0, t_1]\}.$$

Напомним, что краевые условия имели вид $x(t_0) = x(t_1) = 0$. Поэтому естественно рассматривать функционал $\mathcal{L}$ на пространстве

$$\mathring{W}_2^1[t_0, t_1] = \{x \in W_2^1[t_0, t_1] : x(t_0) = x(t_1) = 0\}.$$

Итак, получаем задачу

$$\int\limits_{t_0}^{t_1} (\dot{x}^2 - 2V(t)x\dot{x} - \lambda x^2)\, dt \to \inf, \quad x \in \mathring{W}_2^1[t_0, t_1].$$

Так же, как и раньше, для $x,\ h \in \mathring{W}_2^1[t_0, t_1]$ рассмотрим функцию $\varphi(\mu) = \mathcal{L}(x + \mu h)$ и выпишем $\varphi'(0)$. Имеем:

$$\mathcal{L}(x + \mu h) - \mathcal{L}(x) =$$

$$= \int\limits_{t_0}^{t_1} (2\mu\dot{x}\dot{h} + \mu^2\dot{h}^2 - 2\mu V(t)h\dot{x} - 2\mu V(t)x\dot{h} - 2\mu^2 V(t)h\dot{h} - 2\mu\lambda xh - \mu^2\lambda h^2)\, dt.$$

Отсюда

$$\varphi'(0) = \int\limits_{t_0}^{t_1} (2\dot{x}\dot{h} - 2V(t)h\dot{x} - 2V(t)x\dot{h} - 2\lambda xh)\, dt.$$

Получаем: $\varphi'(0) = 0$ тогда и только тогда, когда

$$\int\limits_{t_0}^{t_1} (\dot{x}\dot{h} - V(t)h\dot{x} - V(t)x\dot{h} - \lambda xh)\, dt = 0, \quad h \in \mathring{W}_2^1[t_0, t_1].$$

Рассмотрим коэффициент при h. Положим $B(t) = \int\limits_{t_0}^{t} (V(s)\dot{x}(s) + \lambda x(s))\,ds$ и проинтегрируем по частям. Получим:

$$\int\limits_{t_0}^{t_1} (\dot{x} + B(t) - V(t)x)\dot{h}\,dt = 0, \quad h \in \mathring{W}_2^1[t_0,\, t_1].$$

В частности, выполнены условия леммы Дюбуа–Реймона; по следствию из нее, $\dot{x} + B(t) - V(t)x = \mathrm{const}$. В частности, $\dot{x} - V(t)x \in AC[t_0,\, t_1]$ и

$$\frac{d}{dt}(\dot{x} - V(t)x) = -V(t)\dot{x} - \lambda x.$$

Это и будет уравнение Штурма–Лиувилля для сингулярного потенциала. Если сделать обозначение $x^{[1]} = \dot{x} - V(t)x$, то последнее уравнение перепишется в виде

$$-\frac{d}{dt}x^{[1]} - V(t)x^{[1]} - V^2(t)x = \lambda x$$

(в такой форме уравнение было в работе Шкаликова–Савчука).

2 Законы сохранения

Снова рассмотрим простейшую задачу вариационного исчисления

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\,dt \to \inf, \quad x(t_0) = x_{(0)},\ x(t_1) = x_{(1)}.$$

Закон сохранения импульса. Пусть L явно не зависит от x_j, то есть

$$L = L(t,\, x_1,\, \ldots,\, x_{j-1},\, x_{j+1},\, \ldots,\, x_n,\, \dot{x}_1,\, \ldots,\, \dot{x}_n).$$

Тогда для j-й переменной уравнение Эйлера имеет вид $-\frac{d}{dt}L_{\dot{x}_j} = 0$. Значит, $L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) = \mathrm{const}$.

В механике $L_{\dot{x}}$ называется обобщенным импульсом, поэтому последнее равенство называется законом сохранения импульса.

Закон сохранения энергии. Пусть L явно не зависит от t, то есть $L = L(x,\, \dot{x})$. Пусть $\hat{x}$ — решение уравнения Эйлера, при этом $\hat{x} \in C^2[t_0,\, t_1]$. Тогда

$$\dot{\hat{x}}(t) L_{\dot{x}}(\hat{x}(t),\, \dot{\hat{x}}(t)) - L(\hat{x}(t),\, \dot{\hat{x}}(t)) = \mathrm{const}. \tag{5}$$

В самом деле, продифференцируем левую часть (5) по t, воспользуемся уравнением Эйлера и получим

$$\frac{d}{dt}(\dot{\hat{x}}(t) L_{\dot{x}}(\hat{x}(t),\, \dot{\hat{x}}(t)) - L(\hat{x}(t),\, \dot{\hat{x}}(t))) =$$

$$= \ddot{\hat{x}}(t) L_{\dot{x}}(\hat{x}(t),\, \dot{\hat{x}}(t)) + \dot{\hat{x}}(t) \frac{d}{dt} L_{\dot{x}}(\hat{x}(t),\, \dot{\hat{x}}(t)) -$$
$$- L_x(\hat{x}(t),\, \dot{\hat{x}}(t))) \dot{\hat{x}}(t) - L_{\dot{x}}(\hat{x}(t),\, \dot{\hat{x}}(t))) \ddot{\hat{x}}(t) = 0.$$

В механике $\dot{x} L_{\dot{x}} - L$ имеет смысл полной энергии системы, поэтому (5) называется законом сохранения энергии.

Достаточное условие C^2-гладкости экстремали у нас уже было (см. предложение 3).

Будет ли в одномерном случае решение уравнения $\dot{x} L_{\dot{x}} - L = c$ экстремалью?

Предложение 4. *Пусть $n = 1$, L явно не зависит от t и дважды непрерывно дифференцируема. Тогда, если функция x не равна константе ни на каком интервале, дважды непрерывно дифференцируема и $\dot{x}(t) L_{\dot{x}}(x(t),\, \dot{x}(t)) - L(x(t),\, \dot{x}(t)) \equiv \mathrm{const}$, то x удовлетворяет уравнению Эйлера.*

Доказательство. Продифференцируем тождество

$$\dot{x}(t) L_{\dot{x}}(x(t),\, \dot{x}(t)) - L(x(t),\, \dot{x}(t)) \equiv \mathrm{const}$$

по t:

$$\ddot{x}(t) L_{\dot{x}}(x(t),\, \dot{x}(t)) + \dot{x}(t) \frac{d}{dt} L_{\dot{x}}(x(t),\, \dot{x}(t)) -$$
$$- L_x(x(t),\, \dot{x}(t)) \dot{x}(t) - L_{\dot{x}}(x(t),\, \dot{x}(t)) \ddot{x}(t) = 0,$$

т.е.

$$\dot{x}(t) \left(\frac{d}{dt} L_{\dot{x}}(x(t),\, \dot{x}(t)) - L_x(x(t),\, \dot{x}(t)) \right) = 0.$$

Так как $\{t \in [t_0, t_1] : \dot{x}(t) \neq 0\}$ всюду плотно, то

$$\frac{d}{dt} L_{\dot{x}}(x(t), \dot{x}(t)) - L_x(x(t), \dot{x}(t)) = 0$$

на всюду плотном множестве. Так как $L \in C^2$, то $L_x \in C^1$, $L_{\dot{x}} \in C^1$; так как $x \in C^2$, то $t \mapsto L_{\dot{x}}(x(t), \dot{x}(t))$, $t \mapsto L_x(x(t), \dot{x}(t))$ непрерывно дифференцируемы. Значит, $t \mapsto \frac{d}{dt} L_{\dot{x}}(x(t), \dot{x}(t)) - L_x(x(t), \dot{x}(t))$ непрерывно и равно нулю на всюду плотном множестве. Отсюда следует уравнение Эйлера–Лагранжа. $\square$

Задача 4. (задача о геодезических на плоскости Лобачевского.) Найти допустимые экстремали в задаче

$$\int\limits_{t_0}^{t_1} \frac{\sqrt{1+\dot{x}^2}}{x}\, dt \to \mathrm{extr}, \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad x > 0.$$

Задача 5. (задача о брахистохроне.) Найти допустимые экстремали в задаче

$$\int\limits_{t_0}^{t_1} \frac{\sqrt{1+\dot{x}^2}}{\sqrt{x}}\, dt \to \mathrm{extr}, \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad x > 0.$$

Задача 6. (задача о минимальной поверхности вращения.) Найти допустимые экстремали в задаче

$$\int\limits_{-T_0}^{T_0} x\sqrt{1+\dot{x}^2}\, dt \to \mathrm{extr}, \quad x(-T_0) = x(T_0) = \xi, \quad x > 0$$

(здесь $\xi > 0$). В зависимости от $\xi > 0$ установить, сколько может быть допустимых экстремалей.

3 Задача Больца

Пусть $L : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ удовлетворяет тем же условиям гладкости, что и в простейшей задаче вариационного исчисления; $l : \mathbb{R}^n \times$

$\mathbb{R}^n \to \mathbb{R}$ непрерывно дифференцируема. Рассмотрим функционал $\mathcal{L}$: $C^1([t_0,\, t_1],\, \mathbb{R}^n) \to \mathbb{R}$, заданный по формуле

$$\mathcal{L}(x) = \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt + l(x(t_0),\, x(t_1)).$$

Производные функции $l(\xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_n)$ по ξ_j будем обозначать $\frac{\partial l}{\partial x_j(t_0)}$, а производные по η_j — через $\frac{\partial l}{\partial x_j(t_1)}$; для $i = 0,\, 1$ через $\frac{\partial l}{\partial x(t_i)}$ будем обозначать n-мерный вектор из $\frac{\partial l}{\partial x_j(t_1)}$, $1 \le j \le n$.

Рассмотрим задачу Больца:

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt + l(x(t_0),\, x(t_1)) \to \inf\,. \tag{6}$$

Функция $\hat{x}$ называется точкой локального минимума в задаче (6), если существует такое $\varepsilon > 0$, что для любой функции $x \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$ такой, что $\|x - \hat{x}\|_{C^1} < \varepsilon$, выполнено $\mathcal{L}(x) \ge \mathcal{L}(\hat{x})$.

Пусть $\hat{x}$ — точка локального минимума в задаче Больца, $h \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$. Так же, как в простейшей задаче вариационного исчисления, определим функцию $\varphi(\lambda) = \mathcal{L}(\hat{x} + \lambda h)$. Тогда $\varphi'(0) = 0$, т.е.

$$\begin{gathered}\int\limits_{t_0}^{t_1} \Big(L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))\dot{h}(t) + L_x(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))h(t) \Big)\, dt + \\ + \tfrac{\partial l}{\partial x(t_0)} h(t_0) + \tfrac{\partial l}{\partial x(t_1)} h(t_1) = 0.\end{gathered} \tag{7}$$

для любого $h \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$.

Предложение 5. *Условие* (7) *эквивалентно выполнению уравнений Эйлера – Лагранжа* (3) *и условий трансверсальности*

$$L_{\dot{x}_j}(t_0,\, \hat{x}(t_0),\, \dot{\hat{x}}(t_0)) = \frac{\partial l}{\partial x_j(t_0)}(\hat{x}(t_0),\, \hat{x}(t_1)),$$

$$L_{\dot{x}_j}(t_1,\, \hat{x}(t_1),\, \dot{\hat{x}}(t_1)) = -\frac{\partial l}{\partial x_j(t_1)}(\hat{x}(t_0),\, \hat{x}(t_1)), \quad 1 \le j \le n;$$

в векторном виде условия трансверсальности записываются как

$$L_{\dot{x}}(t_0,\, \hat{x}(t_0),\, \dot{\hat{x}}(t_0)) = \frac{\partial l}{\partial x(t_0)}(\hat{x}(t_0),\, \hat{x}(t_1)),$$

$$L_{\dot{x}}(t_1,\,\hat{x}(t_1),\,\dot{\hat{x}}(t_1)) = -\frac{\partial l}{\partial x(t_1)}(\hat{x}(t_0),\,\hat{x}(t_1)).$$

Доказательство. Пусть выполнено условие (7). Подставив произвольную функцию $h \in C^1_{0,0}([t_0,\,t_1],\,\mathbb{R}^n)$, получим

$$\int\limits_{t_0}^{t_1} \Big(L_{\dot{x}}(t,\,\hat{x}(t),\,\dot{\hat{x}}(t))\dot{h}(t) + L_x(t,\,\hat{x}(t),\,\dot{\hat{x}}(t))h(t)\Big)\,dt = 0.$$

Ранее было доказано, что функция $t \mapsto L_{\dot{x}}(t,\,\hat{x}(t),\,\dot{\hat{x}}(t))$ абсолютно непрерывна и выполнены уравнения Эйлера – Лагранжа.

Теперь докажем условия трансверсальности. Интегрируя по частям левую часть (7), получаем, что

$$\sum_{j=1}^{n} L_{\dot{x}_j}(t,\,\hat{x}(t),\,\dot{\hat{x}}(t))h_j(t)\big|_{t_0}^{t_1} + \sum_{j=1}^{n} \frac{\partial l}{\partial x_j(t_0)}h_j(t_0) + \sum_{j=1}^{n} \frac{\partial l}{\partial x_j(t_1)}h_j(t_1) = 0$$

для любой функции $h \in C^1([t_0,\,t_1],\,\mathbb{R}^n)$. Фиксируем индекс j и выберем функцию $h \in C^1([t_0,\,t_1],\,\mathbb{R}^n)$ такую, что $h_i(t_0) = h_i(t_1) = 0$ при $i \neq j$, $h_j(t_0) = 1$, $h_j(t_1) = 0$. Тогда получим первое условие трансверсальности. Если взять функцию $h \in C^1([t_0,\,t_1],\,\mathbb{R}^n)$ такую, что $h_i(t_0) = h_i(t_1) = 0$ при $i \neq j$, $h_j(t_0) = 0$, $h_j(t_1) = 1$, то получим второе условие трансверсальности.

Обратно, пусть выполнены уравнения Эйлера и условие трансверсальности. Тогда

$$\int\limits_{t_0}^{t_1} \left(-\frac{d}{dt}L_{\dot{x}}(t,\,\hat{x}(t),\,\dot{\hat{x}}(t)) + L_x(t,\,\hat{x}(t),\,\dot{\hat{x}}(t))\right) h(t)\,dt+$$

$$+\left(-L_{\dot{x}}(t_0,\,\hat{x}(t_0),\,\dot{\hat{x}}(t_0)) + \frac{\partial l}{\partial x(t_0)}(\hat{x}(t_0),\,\hat{x}(t_1))\right) h(t_0)+$$

$$+\left(L_{\dot{x}}(t_1,\,\hat{x}(t_1),\,\dot{\hat{x}}(t_1)) + \frac{\partial l}{\partial x(t_1)}(\hat{x}(t_0),\,\hat{x}(t_1))\right) h(t_1) = 0.$$

Проинтегрировав по частям, получаем (7). □

Отсюда получаем следующий результат.

Теорема 2. *Пусть $\hat{x}$ — точка локального минимума в задаче Больца. Тогда выполнены уравнения Эйлера – Лагранжа и условия трансверсальности.*

Функция $\hat{x} \in C^1([t_0, t_1], \mathbb{R}^n)$, удовлетворяющая уравнениям Эйлера – Лагранжа и условиям трансверсальности, называется допустимой экстремалью.

4 Дифференцируемость в разных смыслах

Пусть X, Y — нормированные пространства, $U \subset X$ — открытое множество, $F : U \to Y$, $x_0 \in U$.

Определение. Отображение F имеет вариацию по Лагранжу в точке x_0, если для любого $h \in X$ существует предел

$$\lim_{\lambda \to 0} \frac{F(x_0 + \lambda h) - F(x_0)}{\lambda} =: F'(x_0)[h].$$

Замечание. Предел двусторонний, λ может быть любого знака; например, функция $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$, в точке $x_0 = 0$ вариации по Лагранжу не имеет.

Если отображение F имеет вариацию по Лагранжу, то определено отображение $F'(x_0) : X \to Y$, заданное по формуле $h \mapsto F'(x_0)[h]$.

Определение. Отображение F дифференцируемо по Гато в точке x_0, если оно имеет вариацию по Лагранжу в т. x_0, а отображение $F'(x_0)$ линейно и непрерывно.

Задача 7. 1) Пусть $F : \mathbb{R}^2 \to \mathbb{R}$ задано равенством $F(x_1, x_2) = \sqrt[3]{x_1^2 x_2}$. Показать, что F имеет вариацию по Лагранжу, но не дифференцируемо по Гато в нуле. 2) Пусть X — бесконечномерное нормированное пространство, $F : X \to \mathbb{R}$ — линейный неограниченный функционал. Показать, что F имеет вариацию по Лагранжу в нуле, но не дифференцируемо по Гато.

Пример. Пусть $F : \mathbb{R}^n \to \mathbb{R}^m$ дифференцируемо по Гато в точке x_0. Тогда $F'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ задается матрицей $A = (a_{i,j})_{1 \le i \le n, 1 \le j \le m}$, где $a_{i,j} = \frac{\partial F_j}{\partial x_i}(x_0)$, т.е. A — матрица Якоби.

Пример. Пусть $\mathcal{L} : C^1([t_0,\, t_1],\, \mathbb{R}^n) \to \mathbb{R}$,

$$\mathcal{L}(x) = \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt + l(x(t_0),\, x(t_1)), \tag{8}$$

где L и l — как в задаче Больца, т.е. $l \in C^1(\mathbb{R}^{2n})$, L, L_x и $L_{\dot{x}}$ непрерывны. Тогда

$$\begin{gathered}\mathcal{L}'(x)[h] = \int\limits_{t_0}^{t_1} (L_x(t,\, x(t),\, \dot{x}(t))h(t) + L_{\dot{x}}(t,\, x(t),\, \dot{x}(t))\dot{h}(t))\, dt + \\ + \tfrac{\partial l}{\partial x(t_0)} h(t_0) + \tfrac{\partial l}{\partial x(t_1)} h(t_1).\end{gathered} \tag{9}$$

Из формулы видно, что отображение $\mathcal{L}$ дифференцируемо по Гато в каждой точке. Из доказанных ранее теорем получаем:

1. в простейшей задаче вариационного исчисления условие $\mathcal{L}'(\hat{x})[h] = 0$ для любого $h \in C^1_{0,0}([t_0,\, t_1],\, \mathbb{R}^n)$ равносильно выполнению уравнений Эйлера–Лагранжа;

2. в задаче Больца условие $\mathcal{L}'(\hat{x})[h] = 0$ для любого $h \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$ равносильно выполнению уравнений Эйлера–Лагранжа и условий трансверсальности.

Определение. Отображение F дифференцируемо по Фреше в точке x_0, если существует линейный непрерывный оператор $A : X \to Y$ такой, что $F(x_0 + h) = F(x_0) + Ah + r(h)$, где отображение r таково, что $\lim\limits_{\|h\| \to 0} \frac{\|r(h)\|_Y}{\|h\|_X} = 0$ (т.е. $r(h) = o(\|h\|)$, $\|h\| \to 0$).

Замечание. 1) Если отображение дифференцируемо по Фреше в т. x_0, то оно дифференцируемо по Гато в т. x_0, при этом $A = F'(x_0)$. 2) Если отображение дифференцируемо по Фреше в т. x_0, то оно непрерывно в т. x_0.

Задача 8. Пусть

$$M = \{(x_1,\, x_2) \in \mathbb{R}^2 :\ x_1 > 0,\ x_2 = x_1^2\},$$

$f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x_1,\, x_2) = \begin{cases} 1, & (x_1,\, x_2) \in M, \\ 0, & (x_1,\, x_2) \notin M. \end{cases}$$

Показать, что f дифференцируемо по Гато, но не дифференцируемо по Фреше в т. $(0, 0)$.

Правила дифференцирования.

Предложение 6. (правило Лейбница). *Пусть $F : U \to Y$, $g : U \to \mathbb{R}$ дифференцируемы по Гато (Фреше) в точке x_0. Тогда отображение $G(x) = g(x)F(x)$ дифференцируемо по Гато (соответственно Фреше) в точке x_0 и $G'(x_0)[h] = g'(x_0)[h]F(x_0) + g(x_0)F'(x_0)[h]$.*

Доказательство. Пусть отображения дифференцируемы по Гато. Имеем

$$\frac{G(x_0+\lambda h) - G(x_0)}{\lambda} = \frac{g(x_0+\lambda h)F(x_0+\lambda h) - g(x_0)F(x_0)}{\lambda} =$$

$$= \frac{(g(x_0) + \lambda g'(x_0)[h] + o(\lambda))(F(x_0) + \lambda F'(x_0)[h] + o(\lambda)) - g(x_0)F(x_0)}{\lambda} \underset{\lambda\to 0}{\to}$$

$$\to g(x_0)F'(x_0)[h] + g'(x_0)[h]F(x_0).$$

Нетрудно видеть, что отображение $h \mapsto g(x_0)F'(x_0)[h] + g'(x_0)[h]F(x_0)$ линейно и непрерывно.

Пусть отображения дифференцируемы по Фреше. Тогда при $\|h\| \to 0$

$$G(x_0+h) = g(x_0+h)F(x_0+h) =$$

$$= (g(x_0) + g'(x_0)[h] + o(\|h\|))(F(x_0) + F'(x_0)[h] + o(\|h\|)) =$$
$$= g(x_0)F(x_0) + g'(x_0)[h]F(x_0) + g(x_0)F'(x_0)[h] + o(\|h\|).$$

□

Предложение 7. (производная композиции). *Пусть X, Y, Z — нормированные пространства, $U \subset X$, $V \subset Y$ — открытые множества, $F : U \to V$, $G : V \to Z$, F дифференцируемо по Гато (Фреше) в т. $x_0 \in U$, G дифференцируемо по Фреше в т. $y_0 = F(x_0)$. Тогда отображение $G \circ F : U \to Z$ дифференцируемо по Гато (Фреше) в т. x_0, при этом*

$$(G\circ F)'(x_0)[h] = G'(y_0)[F'(x_0)[h]], \quad \text{т.е.} \quad (G\circ F)'(x_0) = G'(y_0)\circ F'(x_0).$$

Доказательство. Пусть F дифференцируемо по Гато. Тогда

$$\frac{(G\circ F)(x_0+\lambda h)-(G\circ F)(x_0)}{\lambda}=\frac{G(F(x_0)+\lambda F'(x_0)[h]+o(\lambda))-G(F(x_0))}{\lambda}=$$

$$=\frac{G(F(x_0))+G'(y_0)[\lambda F'(x_0)[h]+o(\lambda)]+r(\lambda F'(x_0)[h]+o(\lambda))-G(F(x_0))}{\lambda}=:S(\lambda);$$

здесь $\frac{r(w)}{\|w\|}\to 0$ при $\|w\|\to 0$. Значит, $r(\lambda F'(x_0)[h]+o(\lambda))=o(\lambda)$ при $\lambda\to 0$. Поэтому

$$S(\lambda)=\frac{\lambda G'(y_0)[F'(x_0)[h]]+o(\lambda)}{\lambda}\underset{\lambda\to 0}{\to}G'(y_0)[F'(x_0)[h]].$$

Пусть F дифференцируемо по Фреше. Тогда

$$(G\circ F)(x_0+h)=G(F(x_0)+F'(x_0)[h]+o(\|h\|))=$$

$$=G(F(x_0))+G'(y_0)[F'(x_0)[h]+o(\|h\|)]+r(F'(x_0)[h]+o(\|h\|))=S(h),$$

где $\frac{r(w)}{\|w\|}\to 0$ при $\|w\|\to 0$. Тогда $r(F'(x_0)[h]+o(\|h\|))=o(\|h\|)$ и

$$S(h)=G(F(x_0))+G'(y_0)[F'(x_0)[h]]+o(\|h\|).$$

□

Задача 9. Построить пример отображений $F:\mathbb{R}\to\mathbb{R}^2$, $G:\mathbb{R}^2\to\mathbb{R}$ таких, что F дифференцируемо по Фреше в т. 0, G дифференцируемо по Гато в т. $(0,\,0)$, $F(0)=(0,\,0)$, при этом $G\circ F$ не имеет вариации по Лагранжу в т. 0. (Можно использовать пример из предыдущей задачи.)

Пусть X, Y — нормированные пространства, $U\subset X$ — открытое множество, $F:U\to Y$, $x_0\in U$. Отображение F называется строго дифференцируемым в точке x_0, если существует линейный непрерывный оператор $A\in L(X,\,Y)$ такой, что для любого $\varepsilon>0$ существует $\delta>0$ такое, что для любых $x_1,\,x_2\in O_\delta(x_0)$ выполнено

$$\|F(x_1)-F(x_2)-A(x_1-x_2)\|_Y\le\varepsilon\|x_1-x_2\|_X.$$

Если подставить $x_1=x_0+h$ и $x_2=x_0$, то получим, что из строгой дифференцируемости следует дифференцируемость по Фреше, при этом $A=F'(x_0)$. Обратное неверно.

Определение строгой дифференцируемости можно переформулировать следующим образом: существует оператор $A \in L(X, Y)$ такой, что константа Липшица отображения $x \mapsto F(x) - Ax$ на $O_\delta(x_0)$ стремится к 0 при $\delta \to 0+$. (Константой Липшица отображения $G : M \to Y$ называется $\sup_{x,\, y \in M} \frac{\|G(x)-G(y)\|_Y}{\|x-y\|_X}$; здесь $M \subset X$.)

Задача 10. Привести пример функции $f : \mathbb{R} \to \mathbb{R}$, всюду дифференцируемой по Фреше, но не строго дифференцируемой в нуле.

Пусть $U \subset X$ — открытое множество, $O(x_0) \subset U$ — окрестность точки x_0, $F : U \to Y$ дифференцируемо по Гато в каждой точке множества $O(x_0)$. Тогда определено отображение $F' : O(x_0) \to L(X, Y)$, заданное по формуле $F' : x \mapsto F'(x)$.

Замечание. Есть три различных объекта: $F'(x_0)[h]$ — это элемент пространства Y; $F'(x_0)$ — это элемент пространства $L(X, Y)$; F' — это отображение из $O(x_0)$ в $L(X, Y)$.

Если отображение $F' : O(x_0) \to L(X, Y)$ непрерывно в т. x_0, то отображение F называется непрерывно дифференцируемым в т. x_0.

Пример. Пусть $A : X \to Y$ — линейный непрерывный оператор. Тогда $A'(x)[h] = A(h)$. Значит, для любого $x \in X$ выполнено $A'(x) = A$, так что A непрерывно дифференцируем в каждой точке.

Пример. Отображение, заданное формулой (8), непрерывно дифференцируемо в каждой точке. Это следует из формулы (9), определения нормы на пространстве $C^1([t_0,\, t_1], \mathbb{R}^n)$ и равномерной непрерывности непрерывной функции на компакте.

Оказывается, из непрерывной дифференцируемости следует строгая дифференцируемость. Для доказательства нам сначала понадобится теорема о среднем.

Напоминание (следствие из теоремы Хана–Банаха). Пусть X — нормированное пространство, $x \in X$. Тогда существует вектор $x^* \in X^*$ такой, что $\|x^*\| = 1$ и $x^*(x) = \|x\|$.

Предложение 8. *(теорема о среднем). Пусть $U \subset X$ — открытое множество, $[x_0,\, x_1] \subset U$, $F : U \to Y$ дифференцируемо по Гато в каждой точке отрезка $[x_0,\, x_1]$. Тогда*

$$\|F(x_1) - F(x_0)\| \leq \sup_{x \in [x_0,\, x_1]} \|F'(x)[x_1 - x_0]\|. \tag{10}$$

Доказательство. Пусть $y^* \in Y^*$, $\|y^*\| = 1$,

$$y^*(F(x_1) - F(x_0)) = \|F(x_1) - F(x_0)\|.$$

Рассмотрим отображение $\varphi : [0,\, 1] \to \mathbb{R}$, $\varphi(t) = y^*(F((1-t)x_0 + tx_1))$. Тогда

$$\varphi'(t) = \lim_{\lambda \to 0} \frac{y^*(F((1-t)x_0 + tx_1 + \lambda(x_1 - x_0))) - y^*(F((1-t)x_0 + tx_1))}{\lambda} =$$

$$= y^*(F'((1-t)x_0 + tx_1)[x_1 - x_0]).$$

По теореме Лагранжа, существует $t \in (0,\, 1)$ такое, что $\varphi(1) - \varphi(0) = \varphi'(t)$. Но $\varphi(1) - \varphi(0) = y^*(F(x_1) - F(x_0)) = \|F(x_1) - F(x_0)\|$. Значит,

$$\|F(x_1) - F(x_0)\| = \varphi'(t) = y^*(F'((1-t)x_0 + tx_1)[x_1 - x_0]) \le$$

$$\le \|F'((1-t)x_0 + tx_1)[x_1 - x_0]\|.$$

□

Задача 11. 1) Если $F : X \to \mathbb{R}$, то существует $x \in [x_0,\, x_1]$ такое, что $F(x_1) - F(x_0) = F'(x)[x_1 - x_0]$. 2) Если $\dim Y > 1$, то утверждение из п. 1 может быть неверным.

Предложение 9. *Если отображение непрерывно дифференцируемо в т. x_0, то оно строго дифференцируемо в т. x_0.*

Доказательство. Нам нужно показать, что для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любых $x_1,\, x_2 \in O_\delta(x_0)$

$$\|F(x_1) - F(x_2) - F'(x_0)[x_1 - x_2]\|_Y \le \varepsilon \|x_1 - x_2\|_X.$$

Рассмотрим отображение $G(x) = F(x) - F'(x_0)[x]$. Тогда, если F дифференцируемо по Гато в т. x, то G дифференцируемо по Гато в т. x и $G'(x)[h] = F'(x)[h] - F'(x_0)[h]$. Значит,

$$\|F(x_1) - F(x_2) - F'(x_0)[x_1 - x_2]\|_Y = \|G(x_1) - G(x_2)\|_Y \overset{(10)}{\le}$$

$$\le \sup_{x \in [x_1,\, x_2]} \|G'(x)[x_1 - x_2]\|_Y \le$$

$$\le \sup_{x \in [x_1,\, x_2]} \|F'(x) - F'(x_0)\|_{L(X,\, Y)} \|x_1 - x_2\|_X.$$

Так как F' непрерывно в x_0 и $x \in [x_1,\, x_2] \subset O_\delta(x_0)$, то при малых $\delta > 0$

$$\sup_{x \in [x_1,\, x_2]} \|F'(x) - F'(x_0)\|_{L(X,Y)} \|x_1 - x_2\|_X \le \varepsilon \|x_1 - x_2\|_X.$$

□

Из строгой дифференцируемости непрерывная дифференцируемость не следует, так как отображение может не быть дифференцируемым по Гато в окрестности точки x_0. Но если дополнительно потребовать дифференцируемость в окрестности, то определения совпадут.

Задача 12. Показать, что если отображение F строго дифференцируемо в точке x_0 и дифференцируемо по Гато в окрестности x_0, то оно непрерывно дифференцируемо в x_0.

5 Оператор Немыцкого

Пусть $\varphi : [t_0,\, t_1] \times \mathbb{R}^n \to \mathbb{R}^m$ непрерывна. Рассмотрим отображение $N_\varphi : C([t_0,\, t_1],\, \mathbb{R}^n) \to C([t_0,\, t_1],\, \mathbb{R}^m)$, заданное по формуле $N_\varphi(x)(t) = \varphi(t,\, x(t))$ (по теореме о непрерывности сложной функции, $N_\varphi(x)(\cdot)$ в самом деле непрерывна). Это отображение называется оператором Немыцкого.

Через $\varphi'_\xi(t,\, \xi)$ обозначим матрицу из частных производных

$$\left(\frac{\partial \varphi_j}{\partial \xi_i}(t,\, \xi)\right)_{1 \le i \le n,\, 1 \le j \le m}.$$

Теорема 3. *Пусть для любого $t \in [t_0,\, t_1]$ отображение $\varphi(t,\, \cdot) : \mathbb{R}^n \to \mathbb{R}^m$ дифференцируемо в каждой точке, при этом φ'_ξ непрерывна на $[t_0,\, t_1] \times \mathbb{R}^n$. Тогда оператор N_φ непрерывно дифференцируем и его производная задается по формуле*

$$N'_\varphi(x)[h](t) = \varphi'_\xi(t,\, x(t))h(t); \tag{11}$$

то есть мы для отображения $\xi \mapsto \varphi(t,\, \xi)$ вычисляем его матрицу Якоби

$$\varphi'_\xi(t,\, \xi) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial \xi_1}(t,\, \xi) & \dots & \frac{\partial \varphi_1}{\partial \xi_n}(t,\, \xi) \\ \dots & \dots & \dots \\ \frac{\partial \varphi_m}{\partial \xi_1}(t,\, \xi) & \dots & \frac{\partial \varphi_m}{\partial \xi_n}(t,\, \xi) \end{pmatrix},$$

затем в качестве $\xi \in \mathbb{R}^n$ подставляем вектор $x(t) \in \mathbb{R}^n$, а после этого действуем полученной матрицей на вектор $h(t) = (h_1(t), \dots,\, h_n(t))^T$.

Доказательство. Сначала покажем, что N_φ имеет вариацию по Лагранжу, имеющую вид (11). Для этого нужно показать, что

$$\|\varphi(\cdot,\, x(\cdot)+\lambda h(\cdot)) - \varphi(\cdot,\, x(\cdot)) - \lambda\varphi'_\xi(\cdot,\, x(\cdot))h(\cdot)\|_C = o(\lambda), \quad \lambda\to 0.$$

Достаточно показать, что для любого $j\in\overline{1,\,m}$ выполнено

$$\|\varphi_j(\cdot,\, x(\cdot)+\lambda h(\cdot)) - \varphi_j(\cdot,\, x(\cdot)) - \lambda\varphi'_{j,\xi}(\cdot,\, x(\cdot))h(\cdot)\|_C = o(\lambda), \quad \lambda\to 0,$$

т.е.

$$\max_{t\in[t_0,\,t_1]} |\varphi_j(t,\, x(t)+\lambda h(t)) - \varphi_j(t,\, x(t)) - \lambda\varphi'_{j,\xi}(t,\, x(t))h(t)| = o(\lambda), \quad \lambda\to 0.$$

Фиксируем $t\in[t_0,\,t_1]$. По теореме Лагранжа, существует $\theta_{\lambda,t}\in[0,\,1]$ такое, что

$$\varphi_j(t,\, x(t)+\lambda h(t)) - \varphi_j(t,\, x(t)) = \lambda\varphi'_{j,\xi}(t,\, x(t)+\theta_{\lambda,t}\lambda h(t))h(t).$$

Значит,

$$|\varphi_j(t,\, x(t)+\lambda h(t)) - \varphi_j(t,\, x(t)) - \lambda\varphi'_{j,\xi}(t,\, x(t))h(t)| =$$

$$= |\varphi'_{j,\xi}(t,\, x(t)+\theta_{t,\lambda}\lambda h(t))h(t) - \varphi'_{j,\xi}(t,\, x(t))h(t)|\cdot|\lambda|.$$

Нам достаточно показать, что

$$\sup_{t\in[t_0,\,t_1]} \|\varphi'_{j,\xi}(t,\, x(t)+\theta_{t,\lambda}\lambda h(t)) - \varphi'_{j,\xi}(t,\, x(t))\| = o(1), \quad \lambda\to 0.$$

В самом деле, пусть $\delta>0$. Рассмотрим множество

$$K_\delta = \{(t,\,\xi):\ t\in[t_0,\,t_1],\ |\xi - x(t)|\le\delta\}.$$

Множество K_1 компактно, поэтому функция $\varphi'_{j,\xi}(\cdot,\,\cdot)$ на нем равномерно непрерывна. Значит, для любого $\varepsilon>0$ найдется $\delta>0$ такое, что для любого $t\in[t_0,\,t_1]$, $w\in\mathbb{R}^n$ такого, что $|w|\le\delta$ выполнено

$$\|\varphi'_{j,\xi}(t,\, x(t)+w) - \varphi'_{j,\xi}(t,\, x(t))\| < \varepsilon. \tag{12}$$

Остается отметить, что при достаточно малых λ для любого $t\in[t_0,\,t_1]$ выполнено $|\theta_{t,\lambda}\lambda h(t)|\le\delta$.

Итак, вариация по Лагранжу вычислена. Легко видеть, что отображение $N'_\varphi(x) : h \mapsto \varphi'_\xi(\cdot,\, x(\cdot))h(\cdot)$ — линейное отображение. Так как матричнозначная функция $\varphi'_\xi(\cdot,\, x(\cdot))$ непрерывна на $[t_0,\, t_1]$, то она ограничена. Отсюда следует, что оператор $N'_\varphi(x)$ ограничен. Тем самым, мы доказали, что N_φ дифференцируем по Гато.

Теперь проверим непрерывную дифференцируемость. Для этого нужно показать, что для любого $x \in C([t_0,\, t_1],\, \mathbb{R}^n)$ выполнено $\|N'_\varphi(y) - N'_\varphi(x)\| \to 0$ при $\|y - x\|_C \to 0$. Так как

$$(N'_\varphi(y) - N'_\varphi(x))h(\cdot) = (\varphi'_\xi(\cdot,\, y(\cdot)) - \varphi'_\xi(\cdot,\, x(\cdot)))h(\cdot),$$

то это следует из (12). $\square$

Задача 13. Пусть $T : L_2[0,\, 1] \to L_2[0,\, 1]$, $Tx(t) = \sin x(t)$. Показать, что T дифференциуемо по Гато в каждой точке, но нигде не дифференцируемо по Фреше.

6 Гладкая задача на минимум на аффинном подпространстве.

Пусть X — нормированное пространство, $Z_0 \subset X$ — линейное подпространство, $\xi_0 \in X$, $Z = Z_0 + \xi_0$, $U \subset Z$ — открытое множество (относительно топологии Z), $f : U \to \mathbb{R}$. Если $x_0 \in U$, то можно определить вариацию по Лагранжу функции f в точке x_0 по направлениям $h \in Z_0$:

$$f'(x_0)[h] = \lim_{\lambda \to 0} \frac{f(x_0 + \lambda h) - f(x_0)}{\lambda}.$$

Рассмотрим задачу поиска точки минимума функции f на U:

$$f(x) \to \min, \quad x \in U. \tag{13}$$

Точка $\hat{x} \in U$ называется точкой локального минимума функции f, если существует такое $\varepsilon > 0$, что для любого $x \in U$ такого, что $\|x - \hat{x}\| < \varepsilon$, выполнено $f(x) \geq f(\hat{x})$.

Если для любого $x \in U$ выполнено $f(x) \geq f(\hat{x})$, то $\hat{x}$ называется точкой глобального минимума.

Аналогично определяется точка локального (глобального) максимума.

Предложение 10. *Пусть* $\hat{x}$ *— точка локального минимума в задаче* (13), *и пусть* f *имеет вариацию по Лагранжу в точке* $\hat{x}$. *Тогда* $f'(\hat{x})[h] = 0$ *для любого* $h \in Z_0$.

Доказательство. Пусть $h \in Z_0$, $\varphi(t) = f(\hat{x} + th)$, где t принадлежит окрестности нуля. Тогда $\varphi'(0) = f'(\hat{x})[h]$ и $t = 0$ является точкой локального минимума функции φ. По теореме Ферма, $\varphi'(0) = 0$. Значит, $f'(\hat{x})[h] = 0$. □

Определение. Пусть X — линейное пространство, $U \subset X$ — выпуклое множество (то есть $(1-\lambda)x + \lambda y \in U$ для любых $x, y \in U$, $\lambda \in [0, 1]$). Функция $f : U \to \mathbb{R}$ называется выпуклой, если для любых $x, y \in U$, $\lambda \in [0, 1]$ выполнено $f((1-\lambda)x + \lambda y) \le (1-\lambda)f(x) + \lambda f(y)$.

Замечание. Из определения следует, что сумма выпуклых функций выпукла.

Предложение 11. *Пусть* X *— линейное нормированное пространство,* $Z_0 \subset X$ *— линейное подпространство,* $\xi_0 \in X$, $Z = Z_0 + \xi_0$, $U \subset Z$ *— выпуклое открытое подмножество,* $f : U \to \mathbb{R}$ *— выпуклая функция,* $\hat{x} \in U$, $f'(\hat{x})[h] = 0$ *для любого* $h \in Z_0$. *Тогда* $\hat{x}$ *— точка глобального минимума.*

Доказательство. Пусть $\hat{x}$ не является точкой минимума. Тогда существует вектор $h \in Z_0$ такой, что $\hat{x} + h \in U$, $f(\hat{x} + h) < f(\hat{x})$. Значит, для $t \in (0, 1]$ выполнено

$$f(\hat{x}+th)-f(\hat{x}) = f((1-t)\hat{x}+t(\hat{x}+h))-f(\hat{x}) \le (1-t)f(\hat{x})+tf(\hat{x}+h)-f(\hat{x}) =$$

$$= t(f(\hat{x}+h) - f(\hat{x})),$$

откуда

$$\frac{f(\hat{x}+th) - f(\hat{x})}{t} \le f(\hat{x}+h) - f(\hat{x}).$$

Поэтому $f'(\hat{x})[h] \le f(\hat{x}+h) - f(\hat{x}) < 0$. □

Пример 1. В простейшей задаче вариационного исчисления $X = C^1([t_0, t_1], \mathbb{R}^n)$, $Z = \{x \in X : \ x(t_0) = x_{(0)}, \ x(t_1) = x_{(1)}\}$, $Z_0 = \{x \in X : x(t_0) = x(t_1) = 0\}$. Тогда

$$\hat{x} \text{ — точка локального минимума } \Rightarrow \ \mathcal{L}'(\hat{x})[h] = 0 \ \forall h \in Z_0 \ \Leftrightarrow$$

$$\Leftrightarrow \int\limits_{t_0}^{t_1} (L_x(t,\, x(t),\, \dot{x}(t))h(t) + L_{\dot{x}}(t,\, x(t),\, \dot{x}(t))\dot{h}(t))\, dt = 0,\ \forall h \in Z_0 \ \Leftrightarrow$$

$$\Leftrightarrow\ -\frac{d}{dt}L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) + L_x(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) = 0,\ L_{\dot{x}}(\cdot,\, \hat{x}(\cdot),\, \dot{\hat{x}}(\cdot)) \in AC[t_0,\, t_1].$$

Если функционал $\mathcal{L}$ выпуклый на Z, то первая импликация $\Rightarrow$ заменяется на $\Leftrightarrow$; более того, минимум оказывается глобальным. Итак, получается

Предложение 12. *Пусть функционал $\mathcal{L}$ выпуклый на Z, $\hat{x}$ — допустимая экстремаль. Тогда $\hat{x}$ является точкой глобального минимума в задаче* (1).

Предложение 13. *Пусть для любого $t \in [t_0,\, t_1]$ отображение $(\xi,\, \eta) \mapsto L(t,\, \xi,\, \eta)$ выпукло ($\xi \in \mathbb{R}^n$, $\eta \in \mathbb{R}^n$). Тогда функционал $\mathcal{L}$ выпуклый на $C^1([t_0,\, t_1],\, \mathbb{R}^n)$.*

Доказательство. Пусть $x,\, y \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$, $\lambda \in [0,\, 1]$. Тогда

$$\mathcal{L}((1-\lambda)x + \lambda y) = \int\limits_{t_0}^{t_1} L(t,\, (1-\lambda)x(t) + \lambda y(t),\, (1-\lambda)\dot{x}(t) + \lambda\dot{y}(t))\, dt \le$$

$$\le \int\limits_{t_0}^{t_1} [(1-\lambda)L(t,\, x(t),\, \dot{x}(t)) + \lambda L(t,\, y(t),\, \dot{y}(t))]\, dt = (1-\lambda)\mathcal{L}(x) + \lambda\mathcal{L}(y).$$

□

Пример 2. В задаче Больца $X = Z_0 = Z = C^1([t_0,\, t_1],\, \mathbb{R}^n)$. Как и в предыдущем примере, если функционал $\mathcal{L}$ выпуклый, то допустимая экстремаль является точкой глобального минимума.

Достаточное условие выпуклости отображения $x \mapsto \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt$ мы уже нашли. Найдем достаточное условие выпуклости отображения $x \mapsto l(x(t_0),\, x(t_1))$.

Предложение 14. *Пусть отображение $(\xi,\, \eta) \mapsto l(\xi,\, \eta)$ выпукло (ξ, $\eta \in \mathbb{R}^n$). Тогда отображение $x \mapsto l(x(t_0),\, x(t_1))$ выпукло.*

Это непосредственно следует из определения выпуклости.

7 Гладкая задача с ограничениями типа равенств

7.1 Постановка задачи и основной результат

Пусть X, Y — нормированные пространства, $U \subset X$ — открытое множество, $f_0 : U \to \mathbb{R}$, $F : U \to Y$. Рассмотрим задачу

$$\begin{cases} f_0(x) \to \inf, \\ F(x) = 0. \end{cases} \tag{14}$$

Точка $x \in U$ называется допустимой для задачи (14), если $F(x) = 0$. Допустимая точка называется точкой локального минимума, если существует такое $\varepsilon > 0$, что для любого допустимого x такого, что $\|x - \hat{x}\| < \varepsilon$, выполнено $f_0(x) \ge f_0(\hat{x})$.

Если $Y = \mathbb{R}^m$, $F(x) = (f_1(x), \dots, f_m(x))$, то (14) переписывается в виде

$$\begin{cases} f_0(x) \to \inf, \\ f_j(x) = 0, \quad 1 \le j \le m. \end{cases} \tag{15}$$

В случае, когда $X = \mathbb{R}^n$, функции f_j непрерывно дифференцируемые ($0 \le j \le m$), в курсе математического анализа доказывалось следующее утверждение.

Предложение 15. *Пусть $X = \mathbb{R}^n$, $\hat{x}$ — точка локального минимума в (15), $f_j \in C^1(U)$, $0 \le j \le m$, $m \le n$, матрица $\left(\frac{\partial f_j}{\partial x_i}(\hat{x})\right)_{1 \le i \le n,\, 1 \le j \le m}$ имеет максимальный ранг. Тогда существуют $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ такие, что*

$$f_0'(\hat{x}) + \sum_{j=1}^{m} \lambda_j f_j'(\hat{x}) = 0,$$

или $\mathcal{L}'(\hat{x}) = 0$, где $\mathcal{L}(x) = f_0(x) + \sum\limits_{j=1}^{m} \lambda_j f_j(x)$.

Функция $\mathcal{L}$ называется функцией Лагранжа.

Здесь мы получим обобщение этого утверждения.

Функция Лагранжа для задачи (14) записывается в виде

$$\mathcal{L}(x) = \lambda_0 f_0(x) + y^*(F(x)), \tag{16}$$

где $\lambda_0 \in \mathbb{R}$, $y^* \in Y^*$. Пара $(\lambda_0,\, y^*) \in \mathbb{R} \times Y^*$ называется множителями Лагранжа.

Рассмотрим частный случай: $Y = \mathbb{R}^m$. Тогда Y^* изоморфно $\mathbb{R}^m$; а именно, каждый элемент y^* задается в виде $y^*(y_1,\, \dots,\, y_m) = \sum\limits_{j=1}^{m} \lambda_j y_j$, где $\lambda_j \in \mathbb{R}$. Поэтому функция Лагранжа будет иметь вид

$$\mathcal{L}(x) = \sum_{j=0}^{m} \lambda_j f_j(x).$$

Сформулируем основной результат.

Теорема 4. (принцип Лагранжа). *Пусть X, Y — банаховы пространства, $\hat{x}$ — точка локального минимума в* (14), *f_0, F строго дифференцируемы в $\hat{x}$,* $\operatorname{Im} F'(\hat{x})$ *замкнут. Тогда найдется* $(\lambda_0,\, y^*) \in \mathbb{R} \times Y^* \backslash \{(0,\, 0)\}$ *такой, что $\lambda_0 f_0'(\hat{x}) + y^* \circ F'(\hat{x}) = 0$, т.е. $\mathcal{L}'(\hat{x}) = 0$, где функция $\mathcal{L}$ имеет вид* (16). *При этом, если оператор $F'(\hat{x})$ сюръективный, то $\lambda_0 \neq 0$.*

Последнее утверждение ($\lambda_0 \neq 0$) верно, так как иначе $y^* \circ F'(\hat{x})[h] = 0$ для любого $h \in X$; из сюръективности $F'(\hat{x})$ следует, что $y^*(y) = 0$ для любого $y \in Y$, т.е. $(\lambda_0,\, y^*) = (0,\, 0)$ — противоречие.

Задача 14. Пусть $A : l_2 \to l_2$,

$$A(x_1,\, x_2,\, \dots,\, x_n,\, \dots) = (x_1,\, x_2/2,\, \dots,\, x_n/n,\, \dots),$$

$(y_1,\, \dots,\, y_n,\, \dots) \in l_2 \backslash \operatorname{Im} A$ (почему такая точка существует?). Рассмотрим задачу

$$\sum_{n=1}^{\infty} y_n x_n \to \inf, \quad A(x_1,\, x_2,\, \dots,\, x_n,\, \dots) = 0.$$

Какая точка будет точкой минимума в этой задаче? Показать, что для этой задачи принцип Лагранжа неверен. Какое из условий теоремы 4 здесь не выполнено?

Задача 15. Привести пример гладких функций $f_0 : \mathbb{R} \to \mathbb{R}$, $f_1 : \mathbb{R} \to \mathbb{R}$ таких, что в задаче $f_0(x) \to \min$, $f_1(x) = 0$ будет существовать точка локального минимума и в принципе Лагранжа будет $\lambda_0 = 0$ (а с $\lambda_0 \neq 0$ принцип Лагранжа не выполнен).

7.2 Касательные векторы

Пусть X — нормированное пространство, $M \subset X$ — непустое множество, $x \in X$. Всюду далее будем обозначать

$$\mathrm{dist}\,(x,\, M) = \inf_{y \in M} \|x - y\|.$$

Пусть $x_0 \in M$. Вектор $h \in X$ называется касательным к M в точке x_0, если

$$\mathrm{dist}\,(x_0 + th,\, M) = o(t), \quad t \to 0$$

(предел двусторонний, т.е. t может быть как положительным, так и отрицательным). Эквивалентное условие: существуют $\varepsilon > 0$ и отображение $r : (-\varepsilon,\, \varepsilon) \to X$ такие, что $\lim_{t \to 0} \frac{\|r(t)\|_X}{t} = 0$ и $x_0 + th + r(t) \in M$ для любого $t \in (-\varepsilon,\, \varepsilon)$.

Множество касательных векторов к M в точке x_0 обозначим через $T_{x_0}M$.

Пример 1. Пусть $F : \mathbb{R}^n \to \mathbb{R}$ — непрерывно дифференцируемое отображение, $F(x_0) = 0$, $\nabla F(x_0) \neq 0$,

$$M = \{x \in \mathbb{R}^n : \ F(x) = 0\}.$$

Тогда

$$T_{x_0}M = \{h \in \mathbb{R}^n : \ \langle \nabla F(x_0),\, h \rangle = 0\}.$$

(Этот факт должен был доказываться на матанализе и выводиться из теоремы о неявной функции.)

Сформулируем необходимое условие локального минимума функции на множестве в терминах касательного вектора.

Пусть X — нормированное пространство, $U \subset X$ — открытое множество, $M \subset U$ — непустое множество, $f_0 : U \to \mathbb{R}$. Рассмотрим задачу

$$\begin{cases} f_0(x) \to \inf, \\ x \in M. \end{cases} \tag{17}$$

Точка $\hat{x} \in M$ называется точкой локального минимума в задаче (17), если существует $\varepsilon > 0$ такое, что для любого $x \in M$ такого, что $\|x - \hat{x}\| < \varepsilon$ выполнено $f_0(x) \geq f_0(\hat{x})$.

Предложение 16. *Пусть $\hat{x} \in M$ — точка локального минимума в задаче* (17)*, функция f_0 дифференцируема по Фреше в точке $\hat{x}$. Тогда $f_0'(\hat{x})[h] = 0$ для любого $h \in T_{\hat{x}}M$.*

Доказательство. Пусть $h \in T_{\hat{x}}M$. Тогда существуют $\delta > 0$ и функция $r : (-\delta, \delta) \to X$ такие, что $\|r(t)\|_X = o(t)$ при $t \to 0$ и $\hat{x}+th+r(t) \in M$ для любого $t \in (-\delta, \delta)$. Положим $\varphi(t) = f_0(\hat{x} + th + r(t))$. Тогда 0 является точкой локального минимума функции φ. Значит, если φ дифференцируема в нуле, то $\varphi'(0) = 0$ по теореме Ферма.

Имеем: $\varphi(t) = \varphi(0) + f_0'(\hat{x})[th+r(t)] + \omega(th+r(t))$, где $\omega(\xi) = o(\xi)$ при $\xi \to 0$. Значит, $\varphi(t) = \varphi(0) + t \cdot f_0'(\hat{x})[h] + o(t)$ и поэтому $\varphi'(0) = f_0'(\hat{x})[h]$. Тем самым, $f_0'(\hat{x})[h] = 0$. □

Задача 16. Верно ли то же самое утверждение, если f_0 дифференцируема по Гато?

Рассмотрим еще раз множество M из примера 1. Получаем, что если $\hat{x}$ — точка локального минимума функции f_0 на множестве M, то $\langle \nabla f_0(\hat{x}), h\rangle = 0$ для любого вектора $h \in \mathbb{R}^n$ такого, что $\langle \nabla F(\hat{x}), h\rangle = 0$. Это означает, что вектор $\nabla f_0(\hat{x})$ ортогонален гиперплоскости $\{h \in \mathbb{R}^n : \langle \nabla F(\hat{x}), h\rangle = 0\}$, то есть пропорционален $\nabla F(\hat{x})$. Это и означает выполнение принципа Лагранжа для этого частного случая.

Для доказательства принципа Лагранжа в общем случае нам понадобятся теорема о касательном пространстве, теорема об аннуляторе ядра и теорема о нетривиальном аннуляторе.

7.3 Лемма о правом обратном

Будем далее обозначать через B_X замкнутый единичный шар пространства X.

Сформулируем теорему Банаха об открытом отображении (без доказательства).

Теорема 5. *Пусть X, Y — банаховы пространства, $A : X \to Y$ — линейный непрерывный сюръективный оператор. Тогда существует $\delta > 0$ такое, что $A(B_X) \supset \delta \cdot B_Y$.*

Эта теорема доказывается почти так же, как и теорема Банаха об обратном операторе.

Теперь докажем лемму о правом обратном.

Лемма 1. *Пусть X, Y — банаховы пространства, $A : X \to Y$ — линейный непрерывный сюръективный оператор. Тогда существуют отображение $R : Y \to X$ и число $C > 0$ такие, что для любого $y \in Y$ выполнено $ARy = y$ и $\|Ry\| \le C\|y\|$.*

Доказательство. По теореме 5, существует $\delta > 0$ такое, что $\delta B_Y \subset A(B_X)$. Значит, если $\|y\| = \delta$, то $y = Ax(y)$, где $\|x(y)\| \le 1$. Тогда мы можем определить отображение R равенствами $Ry = \frac{\|y\|}{\delta} x\left(\frac{\delta y}{\|y\|}\right)$, $y \neq 0$; $R(0) = 0$. $\square$

7.4 Теорема Люстерника и теорема о касательном пространстве

Докажем теорему Люстерника.

Теорема 6. *Пусть X, Y — банаховы пространства, $U \subset X$ — открытое множество, $x_0 \in U$, $F : U \to Y$ строго дифференцируемо в точке x_0, оператор $F'(x_0)$ сюръективен. Пусть $F(x_0) = 0$, $M = \{x \in U : F(x) = 0\}$. Тогда существуют окрестность V точки x_0 и константа $K > 0$ такие, что для любого $x \in V$ выполнено*

$$\operatorname{dist}(x, M) \le K\|F(x)\|.$$

Доказательство. Нам нужно каждой точке x из достаточно малой окрестности x_0 точку $\varphi(x) \in M$ такую, что

$$\|x - \varphi(x)\| \le K\|F(x)\|. \tag{18}$$

Точка $\varphi(x)$ будет искаться методом, похожим на метод Ньютона. Сначала напомним этот метод для функции одной переменной и затем по аналогии напишем итерационный процесс для поиска $\varphi(x)$.

Итак, пусть $f : \mathbb{R} \to \mathbb{R}$ — непрерывно дифференцируемая функция, $f(\hat{x}) = 0$, $f'(\hat{x}) \neq 0$. Тогда метод Ньютона для поиска $\hat{x}$ записывается следующим образом. Если построена точка x_n, то в ней проводим касательную к графику f и ищем x_{n+1} — точку пересечения касательной с осью абсцисс. Получаем уравнение $f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0$. Отсюда

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Для поиска $\varphi(x)$ напишем такую же формулу, только вместо деления на $f'(x_n)$ будет применение правого обратного оператора к $F'(x_0)$. Напомним: поскольку X, Y банаховы, оператор $F'(x_0)$ сюръективен, то существуют отображение $R : Y \to X$ и константа $C > 0$ такие, что для любого $y \in Y$ выполнено $F'(x_0)Ry = y$ и

$$\|Ry\| \leq C\|y\|. \tag{19}$$

Итак, записываем итерационный процесс:

$$x_1 = x, \quad x_{n+1} = x_n - RF(x_n). \tag{20}$$

Мы покажем, что если x принадлежит достаточно малой окрестности точки x_0, то x_n будет принадлежать U при всех n (и, значит, $F(x_n)$ определено) и $\{x_n\}_{n\in\mathbb{N}}$ будет фундаментальной.

Сначала напишем серию оценок при определенных предположениях, а потом покажем, что если x достаточно близко к x_0, то они будут выполнены для всех n.

Сначала применим к (20) оператор $F'(x_0)$:

$$F'(x_0)(x_{n+1} - x_n) + F(x_n) = 0. \tag{21}$$

Возьмем произвольное $\varepsilon > 0$ и выберем для него $\delta > 0$ в соответствии с определением строгой дифференцируемости. Тогда, если x_n, $x_{n+1} \in O_\delta(x_0)$, то

$$\|F(x_{n+1}) - F(x_n) - F'(x_0)(x_{n+1} - x_n)\| \leq \varepsilon\|x_{n+1} - x_n\|.$$

Учитывая (21), получаем

$$\|F(x_{n+1})\| \leq \varepsilon\|x_{n+1} - x_n\|. \tag{22}$$

Снова из (20) получаем:

$$\|x_{n+1} - x_n\| = \|RF(x_n)\| \leq C\|F(x_n)\| \overset{(22)}{\leq} C\varepsilon\|x_n - x_{n-1}\| \tag{23}$$

(последнее неравенство выполнено при $n \geq 2$, в предположении, что x_{n-1} тоже принадлежит $O_\delta(x_0)$).

Пусть $\varepsilon = \frac{1}{2C}$. Тогда

$$\|x_{n+1} - x_n\| \leq \frac{1}{2}\|x_n - x_{n-1}\|. \tag{24}$$

Значит, если $x_1, \ldots, x_n, \ldots, x_{n+k}$ принадлежат $O_\delta(x_0)$, то

$$\begin{aligned} \|x_{n+k} - x_n\| \le \|x_{n+k} - x_{n+k-1}\| + \cdots + \|x_{n+1} - x_n\| \overset{(24)}{\le} \\ \le (2^{-n-k+2} + \cdots + 2^{-n+1})\|x_2 - x_1\| \le 2^{-n+2}\|x_2 - x_1\|. \end{aligned} \tag{25}$$

Величина $\|x_2-x_1\|$ оценивается с помощью (23): $\|x_2-x_1\| \le C\|F(x_1)\|$; так как отображение F непрерывно в точке x_0 (в силу строгой дифференцируемости), то правая часть будет мала, если $x_1 = x$ будет принадлежать малой окрестности точки x_0. А значит, учитывая (25), можно добиться того, что для любого $n \in \mathbb{N}$ точки x_n будут принадлежать $O_\delta(x_0)$ и все оценки будут верны.

Из (25) следует, что $\{x_n\}_{n\in\mathbb{N}}$ фундаментальна. Так как X полно, то существует точка $\varphi(x) := \lim\limits_{n\to\infty} x_n$; снова применяем (25) и получаем:

$$\|x_n - x_1\| \le 4\|x_2 - x_1\| \overset{(23)}{\le} 4C\|F(x_1)\|;$$

значит, $\|\varphi(x) - x\| \le 4C\|F(x)\|$.

Из (22) и (25) следует, что $F(x_n) \underset{n\to\infty}{\to} 0$; с другой стороны, из непрерывности F в окрестности точки x_0 (которая следует из строгой дифференцируемости) получаем, что $F(x_n) \underset{n\to\infty}{\to} F(\varphi(x))$ (если x в малой окрестности точки x_0). Значит, $F(\varphi(x)) = 0$, т.е. $\varphi(x) \in M$. $\square$

Теперь докажем теорему о касательном пространстве.

Теорема 7. *Пусть X, Y — банаховы пространства, $U \subset X$ — открытое множество, $x_0 \in U$, $F : U \to Y$ строго дифференцируемо в точке x_0, $F(x_0) = 0$, оператор $F'(x_0)$ сюръективен. Пусть $M = \{x \in U : F(x) = 0\}$. Тогда $T_{x_0}M = \ker F'(x_0)$.*

Доказательство. Сначала докажем включение $T_{x_0}M \subset \ker F'(x_0)$. Пусть $h \in T_{x_0}M$. Тогда существуют $\varepsilon > 0$ и функция $r : (-\varepsilon, \varepsilon) \to X$ такие, что $x_0 + th + r(t) \in M$ и $r(t) = o(t)$ при $t \to 0$. Значит, $F(x_0 + th + r(t)) = 0$ для любого $t \in (-\varepsilon, \varepsilon)$.

Отображение F дифференцируемо по Фреше, поэтому

$$F(x_0 + \xi) = F(x_0) + F'(x_0)[\xi] + \omega(\xi), \quad \omega(\xi) = o(\|\xi\|),\ \xi \to 0.$$

Значит, учитывая, что $F(x_0) = 0$, получаем

$$F'(x_0)[th + r(t)] + \omega(th + r(t)) = 0, \quad t \in (-\varepsilon, \varepsilon).$$

Поэтому $tF'(x_0)[h] = o(t)$ при $t \to 0$. Это возможно только если $F'(x_0)[h] = 0$, т.е. $h \in \ker F'(x_0)$.

Теперь докажем, что $\ker F'(x_0) \subset T_{x_0}M$. Пусть $F'(x_0)[h] = 0$. Применяя теорему Люстерника к вектору $x_0 + th$ при малых t, получаем, что существуют $K > 0$ и $\varepsilon > 0$ такие, что для любого $t \in (-\varepsilon, \varepsilon)$ существует точка $r(t) \in X$ такая, что

$$x_0 + th + r(t) \in M, \tag{26}$$

$$\|r(t)\| \le K\|F(x_0 + th)\|. \tag{27}$$

Так как $F(x_0) = 0$ и $F'(x_0)[h] = 0$, то

$$F(x_0 + th) = F(x_0) + F'(x_0)[th] + o(t) = o(t),\ t \to 0.$$

Значит, $\|r(t)\| \overset{(27)}{=} o(t)$, $t \to 0$. В силу (26), это означает, что $h \in T_{x_0}M$. $\square$

Замечание. Включение $T_{x_0}M \subset \ker F'(x_0)$ выполнено всегда, даже если X или Y не полно или оператор $F'(x_0)$ не сюръективен.

7.5 Теоремы отделимости

Пусть X — линейное пространство, $M \subset X$ — выпуклое множество, $0 \in M$. Предположим, что для любого $x \in X$ существует $\lambda > 0$ такое, что $\frac{x}{\lambda} \in M$. Функционалом Минковского множества M называется отображение $p_M : X \to \mathbb{R}_+$, заданное по формуле

$$p_M(x) = \inf\left\{\lambda > 0 :\ \frac{x}{\lambda} \in M\right\}.$$

Пример. Пусть $(X, \|\cdot\|)$ — нормированное пространство, M — единичный шар в X. Тогда

$$p_M(x) = \inf\left\{\lambda > 0 :\ \left\|\frac{x}{\lambda}\right\| \le 1\right\} = \|x\|.$$

Отметим следующие свойства функционала Минковского:

- Пусть $M \subset N$. Тогда $p_M(x) \ge p_N(x)$.
- Пусть $t > 0$. Тогда $p_{tM}(x) = \frac{p_M(x)}{t}$.

Функция $f : X \to \mathbb{R}$ называется положительно-однородной, если для любых $x \in X$, $\lambda \geq 0$ выполнено $f(\lambda x) = \lambda f(x)$.

Предложение 17. *Функционал Минковского выпуклый и положительно-однородный.*

Доказательство. Сначала докажем положительную однородность. Легко видеть, что $p_M(0) = 0$. Пусть $x \in X$, $\lambda > 0$. Тогда

$$p_M(\lambda x) = \inf\left\{t > 0 :\ \frac{\lambda x}{t} \in M\right\} = \inf\left\{\lambda s > 0 :\ \frac{x}{s} \in M\right\} = \lambda p_M(x).$$

Пусть $x_1,\ x_2 \in X$. Покажем, что $p_M(x_1 + x_2) \leq p_M(x_1) + p_M(x_2)$. В самом деле, пусть $t_1 > p_M(x_1)$, $t_2 > p_M(x_2)$. Тогда $\frac{x_1}{t_1} \in M$, $\frac{x_2}{t_2} \in M$. Значит,

$$\frac{x_1 + x_2}{t_1 + t_2} = \frac{t_1}{t_1 + t_2} \cdot \frac{x_1}{t_1} + \frac{t_2}{t_1 + t_2} \cdot \frac{x_2}{t_2} \in M.$$

Поэтому $p_M(x_1 + x_2) \leq t_1 + t_2$; остается сделать предельный переход при $t_1 \to p_M(x_1)$, $t_2 \to p_M(x_2)$.

Теперь легко проверить выпуклость p_M: если $x_1,\ x_2 \in X$, $\lambda \in [0,\,1]$, то

$$p_M((1-\lambda)x_1 + \lambda x_2) \leq p_M((1-\lambda)x_1) + p_M(\lambda x_2) = (1-\lambda)p_M(x_1) + \lambda p_M(x_2).$$

□

Еще раз напомним теорему Хана–Банаха из функционального анализа:

Теорема 8. *Пусть X — линейное пространство, $L \subset X$ — подпространство, $p : X \to \mathbb{R}$ — выпуклый положительно-однородный функционал, $f_0 : L \to \mathbb{R}$ — линейный функционал такой, что $f_0(x) \leq p(x)$ для любого $x \in L$. Тогда существует линейный функционал $f : X \to \mathbb{R}$ такой, что $f|_L = f_0$ и $f(x) \leq p(x)$ для любого $x \in X$.*

Теорема 9. *Пусть X — нормированное пространство, $A,\ B \subset X$ — непустые выпуклые множества,* $\operatorname{int} A \neq \varnothing$, $(\operatorname{int} A) \cap B = \varnothing$. *Тогда существует функционал $x^* \in X^* \backslash \{0\}$ такой, что*

$$\sup_{x \in A} x^*(x) \leq \inf_{x \in B} x^*(x).$$

Геометрический смысл теоремы следующий. Пусть

$$\sup_{x\in A} x^*(x) \le c \le \inf_{x\in B} x^*(x).$$

Рассмотрим гиперплоскость

$$L = \{x \in X :\ x^*(x) = c\}.$$

Она разбивает X на два полупространства: $\Pi_- = \{x \in X :\ x^*(x) \le c\}$ и $\Pi_+ = \{x \in X :\ x^*(x) \ge c\}$. Тогда $A \subset \Pi_-$, $B \subset \Pi_+$, т.е. гиперплоскость L разделяет множества A и B.

Сначала докажем теорему в частном случае.

Лемма 2. *Пусть $M \subset X$ — выпуклое множество,* $0 \in \operatorname{int} M$, $x_0 \notin \operatorname{int} M$. *Тогда существует функционал $x^* \in X^*\setminus\{0\}$ такой, что*

$$\sup_{x\in M} x^*(x) \le x^*(x_0).$$

Доказательство. Так как M выпукло и $0 \in \operatorname{int} M$, то для него корректно определен функционал Минковского.

Покажем, что $p_M(x_0) \ge 1$. В самом деле, пусть $p_M(x_0) < 1$. Тогда существует $\lambda \in (0,\, 1)$ такое, что $\hat{x} := x_0/\lambda \in M$. Далее, существует $\delta > 0$ такое, что $B_\delta(0) \subset M$. Тогда $C = \operatorname{conv}(B_\delta(0) \cup \{\hat{x}\}) \subset M$.

Покажем, что $x_0 \in \operatorname{int} C$ (тогда получим противоречие с тем, что $x_0 \notin \operatorname{int} M$). В самом деле, $x_0 = (1-\lambda)0 + \lambda\hat{x}$ и

$$B_{(1-\lambda)\delta}(x_0) = (1-\lambda)B_\delta(0) + \lambda\hat{x} \subset C.$$

Определим функционал f_0 на $\operatorname{span}\{x_0\}$ по формуле $f_0(tx_0) = t$. Если $t < 0$, то $f_0(tx_0) < 0 \le p_M(tx_0)$, а если $t \ge 0$, то

$$f_0(tx_0) = t \le t \cdot p_M(x_0) = p_M(tx_0).$$

По теореме Хана–Банаха, существует линейный функционал f такой, что $f(x) \le p_M(x)$ для любого $x \in X$ и $f|_{\operatorname{span}\{x_0\}} = f_0$. Тогда 1) $f \ne 0$, 2) $f(x_0) = 1$, 3) для любого $x \in M$ выполнено $f(x) \le p_M(x) \le 1$. Из 2) и 3) следует, что $\sup_{x\in M} f(x) \le f(x_0)$.

Покажем, что f непрерывен. В самом деле, для любого $x \in X$

$$|f(x)| = \max\{f(-x),\, f(x)\} \le \max\{p_M(-x),\, p_M(x)\} \le$$

$$\le \max\{p_{B_\delta(0)}(x),\, p_{B_\delta(0)}(-x)\} = \frac{\|x\|}{\delta}$$

(напомним, что $B_\delta(0) \subset M$, поэтому $p_M(x) \le p_{B_\delta(0)}(x)$, $x \in X$). $\square$

Следствие 1. *Пусть $M \subset X$ — выпуклое множество с непустой внутренностью, $x_0 \notin \operatorname{int} M$. Тогда существует функционал $x^* \in X^* \backslash \{0\}$ такой, что $x^*(x_0) \geq \sup_{x \in M} x^*(x)$.*

Доказательство. Пусть $\hat{x} \in \operatorname{int} M$. Применим предыдущую лемму к множеству $M - \hat{x}$ и точке $x_0 - \hat{x}$ и получим требуемое утверждение. $\square$

Предложение 18. *Пусть A — выпуклое множество, $\operatorname{int} A \neq \varnothing$. Тогда $A \subset \overline{\operatorname{int} A}$.*

Доказательство. Пусть $O_\delta(x_0) \subset A$, $x \in A$. Тогда

$$C = \operatorname{conv}(O_\delta(x_0) \cup \{x\}) \subset A,$$

при этом $(1-\lambda)x_0 + \lambda x \in \operatorname{int} A$ для $0 \leq \lambda < 1$. Значит, $x \in \overline{\operatorname{int} A}$. $\square$

Пусть A, $B \subset X$. Положим

$$A + B = \{x + y : \ x \in A, \, y \in B\}.$$

Заметим, что если множества A, B выпуклы, то $A + B$ тоже выпукло.

Доказательство теоремы 9. Нам нужно найти функционал $x^* \in X^* \backslash \{0\}$ такой, что для любых $x \in A$, $y \in B$ выполнено $x^*(x) \leq x^*(y)$. Это эквивалентно каждому из условий:

1. $x^*(x) \leq x^*(y)$, $x \in \operatorname{int} A$, $y \in B$ (в силу предложения 18 и непрерывности x^*);

2. $x^*(x + (-y)) \leq 0$, $x \in \operatorname{int} A$, $y \in B$;

3. $x^*(z) \leq 0$ для любого $z \in (\operatorname{int} A) + (-B)$.

Множество $(\operatorname{int} A)+(-B)$ выпукло, открыто и не содержит 0 (т.к. $(\operatorname{int} A) \cap B = \varnothing$). Применяя следствие 1, находим требуемый функционал x^*. $\square$

Следующая теорема — о строгой отделимости точки от замкнутого выпуклого множества.

Теорема 10. *Пусть X — нормированное пространство, A — непустое выпуклое замкнутое множество, $x_0 \notin X$. Тогда существует функционал $x^* \in X^* \backslash \{0\}$ такой, что*

$$\sup_{x \in A} x^*(x) < x^*(x_0).$$

Доказательство. Так как A замкнуто, то существует $\varepsilon > 0$ такое, что $B_\varepsilon(x_0) \cap A = \varnothing$. По теореме 9, существует функционал $x^* \in X^* \backslash \{0\}$ такой, что

$$\sup_{x\in A} x^*(x) \le \inf_{x\in B_\varepsilon(x_0)} x^*(x) < x_*(x_0).$$

Теорема доказана. □

7.6 Вид линейного непрерывного функционала на декартовом произведении пространств

Пусть $X_1, \ldots, X_n$ — линейные нормированные пространства. На декартовом произведении $X_1 \times \cdots \times X_n$ вводится норма

$$\|(x_1, \ldots, x_n)\| = \max\{\|x_1\|_{X_1}, \ldots, \|x_n\|_{X_n}\}.$$

Предложение 19. *Пусть $X_1, \ldots, X_n$ — линейные нормированные пространства. Тогда $x^* \in (X_1 \times \cdots \times X_n)^*$ тогда и только тогда, когда существуют функционалы $x_1^* \in X_1^*, \ldots, x_n^* \in X_n^*$ такие, что*

$$x^*(x_1, \ldots, x_n) = \sum_{j=1}^n x_j^*(x_j), \quad (x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n. \tag{28}$$

Доказательство. Пусть $x^* \in (X_1 \times \cdots \times X_n)^*$. Определим функционалы x_j^* равенством $x_j^*(z) = x^*(0, \ldots, 0, z, 0, \ldots, 0)$ ($z \in X_j$ стоит на j-м месте). Тогда выполнено (28), x_j^* линеен и $\|x_j^*\| \le \|x^*\|$.

Обратно, функционал вида (28) линеен и

$$|x^*(x_1, \ldots, x_n)| \le \sum_{j=1}^n |x_j^*(x_j)| \le \left(\sum_{j=1}^n \|x_j^*\|\right) \|(x_1, \ldots, x_n)\|.$$

Значит, x^* непрерывен. □

7.7 Аннулятор ядра оператора

Пусть X — нормированное пространство, $L \subset X$ — подпространство. Тогда аннулятором L называется

$$L^\perp = \{x^* \in X^* : \; \forall x \in L \;\; x^*(x) = 0\}.$$

Напомним определение сопряженного оператора. Пусть $A : X \to Y$ — линейный непрерывный оператор. Тогда сопряженный оператор $A^* : Y^* \to X^*$ определяется равенством $(A^*y^*)(x) = y^*(Ax)$.

Теорема 11. *Пусть X, Y — банаховы пространства, $A : X \to Y$ — сюръективный оператор. Тогда* $(\ker A)^\perp = \operatorname{Im} A^*$.

Доказательство. 1) Пусть $x^* \in \operatorname{Im} A^*$, то есть $x^* = A^*y^*$, где $y^* \in Y^*$. Тогда для любого $x \in \ker A$ выполнено

$$x^*(x) = A^*y^*(x) = y^*(Ax) = 0.$$

2) Пусть $x^* \in (\ker A)^\perp$. Рассмотрим множество

$$M = \{(\mu,\, y) \in \mathbb{R} \times Y :\ \exists x \in X :\ \mu > x^*(x),\, y = Ax\}.$$

Множество M имеет непустую внутренность. В самом деле, по теореме Банаха об открытом отображении, существует $\delta > 0$ такое, что $A(B_X) \supset \delta B_Y$. С другой стороны, если $\mu > \|x^*\|$, то $\mu > x^*(x)$ для любого $x \in B_X$. Значит,

$$M \supset (\|x^*\|,\, +\infty) \times \delta B_Y. \tag{29}$$

Далее, $0 \notin M$; иначе существует $x \in X$ такое, что $0 > x^*(x)$ и $0 = Ax$, т.е. $x^* \notin (\ker A)^\perp$.

По теореме отделимости, существует $z^* \in (\mathbb{R}\times Y)^*$ такой, что $z^*(0,\, 0) \le \inf_{(\mu,\, y)\in M} z^*(\mu,\, y)$. То есть существуют $(\lambda,\, y^*) \in \mathbb{R}\times Y^* \setminus \{(0,\, 0)\}$ такие, что

$$0 \le \lambda\mu + y^*(y),\ (\mu,\, y) \in M.$$

Заметим, что $\lambda \geq 0$; иначе можно в правой части устремить $\mu \to +\infty$ и получить $0 \le -\infty$. Случай $\lambda = 0$ также не подходит. В самом деле, если $\lambda = 0$, то $0 \le y^*(y)$ для $(\mu,\, y) \in M$; в частности, в силу (29), $0 \le y^*(y)$ для $y \in \delta B_Y$. Это возможно только если $y^* = 0$. Но тогда $(\lambda,\, y^*) = (0,\, 0)$.

Итак, $\lambda > 0$. Можно считать, что $\lambda = 1$. Тогда $\mu + y^*(y) \ge 0$ для $(\mu,\, y) \in M$. В частности, для любого $\varepsilon > 0$, $x \in X$

$$x^*(x) + \varepsilon + y^*(Ax) \ge 0.$$

Устремив $\varepsilon \to +0$, получаем

$$x^*(x) + y^*(Ax) \ge 0, \quad x \in X.$$

Заменив x на $-x$, получим

$$x^*(x) + y^*(Ax) = 0, \quad x \in X.$$

Значит, $x^* = -y^* \circ A = A^*(-y^*)$. □

7.8 Завершение доказательства принципа Лагранжа

Пусть выполнены условия теоремы 4.

Сначала рассмотрим случай, когда $F'(\hat{x})$ сюръективно.

Уже было доказано: если $\hat{x}$ — точка локального минимума, то $f_0'(\hat{x})[h] = 0$ для любого $h \in T_{\hat{x}}M$, где $M = \{x : \ F(x) = 0\}$. По теореме о касательном пространстве, $f_0'(\hat{x})[h] = 0$ для любого $h \in \ker F'(\hat{x})$, то есть $f_0'(\hat{x}) \in (\ker F'(\hat{x}))^{\perp}$. По теореме об аннуляторе ядра, $f_0'(\hat{x}) \in \mathrm{Im}\,(F'(\hat{x}))^*$, то есть существует функционал $z^* \in Y^*$ такой, что $f_0'(\hat{x}) = z^* \circ F'(\hat{x})$. Остается положить $y^* = -z^*$, $\lambda_0 = 1$.

Теперь рассмотрим случай, когда $\mathrm{Im}\, F'(x_0) \neq Y$ — замкнутое подпространство. Нам понадобится лемма о нетривиальном аннуляторе.

Лемма 3. *Пусть X — нормированное пространство, $L \subset X$ — замкнутое подпространство, $L \neq X$. Тогда существует $x^* \in L^{\perp} \backslash \{0\}$.*

Доказательство. Пусть $x_0 \in X \backslash L$. По теореме о строгой отделимости, существует функционал $x^* \in X^*$ такой, что $x^*(x_0) > \sup_{x \in L} x^*(x)$. Докажем, что $x^*(x) = 0$ для любого $x \in L$. В самом деле, пусть $x^*(x_1) \neq 0$ для некоторого $x_1 \in L$. Так как $\lambda x_1 \in L$ для любого $\lambda \in \mathbb{R}$, то $\sup_{x \in L} x^*(x) = +\infty$ — противоречие. Значит, $x^* \in L^{\perp}$ и $x^*(x_0) > 0$. $\square$

Итак, пусть $\mathrm{Im}\, F'(x_0) \neq Y$ — замкнутое подпространство. По лемме о нетривиальном аннуляторе, существует функционал $y^* \in Y^* \backslash \{0\}$ такой, что $y^*|_{\mathrm{Im}\, F'(\hat{x})} = 0$, т.е. $y^*(F'(\hat{x})[h]) = 0$ для любого $h \in X$. Это означает, что выполнен принцип Лагранжа с $\lambda_0 = 0$.

Замечание. Если $\mathrm{Im}\, F'(x_0)$ не замкнут, при этом $Z := \overline{\mathrm{Im}\, F'(x_0)} \neq Y$, то принцип Лагранжа тоже верен, так как можно применить лемму о нетривиальном аннуляторе к Z.

8 Изопериметрическая задача

Пусть $n \in \mathbb{N}$, $m \in \mathbb{N}$, $L_j : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ $(0 \leq j \leq m)$ — непрерывные функции. Пусть $c_j \in \mathbb{R}$, $1 \leq j \leq m$, $x_{(0)}$, $x_{(1)} \in \mathbb{R}^n$. Изопериметрическая задача — это задача на экстремум следующего вида:

$$\begin{cases} \int\limits_{t_0}^{t_1} L_0(t,\, x(t),\, \dot{x}(t))\, dt \to \inf, \\ \int\limits_{t_0}^{t_1} L_j(t,\, x(t),\, \dot{x}(t))\, dt = c_j, \quad 1 \le j \le m, \\ x(t_0) = x_{(0)},\ x(t_1) = x_{(1)}, \\ x \in C^1([t_0,\, t_1],\, \mathbb{R}^n). \end{cases} \tag{30}$$

Обозначим $\mathcal{L}_j(x) = \int\limits_{t_0}^{t_1} L_j(t,\, x(t),\, \dot{x}(t))\, dt$, $0 \le j \le m$.

Функция $x \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$ называется допустимой для задачи (30), если $\mathcal{L}_j(x) = c_j$, $1 \le j \le m$, $x(t_0) = x_{(0)}$, $x(t_1) = x_{(1)}$. Допустимая функция $\hat{x}$ называется точкой локального минимума в задаче (30), если существует такое $\varepsilon > 0$, что для любой допустимой функции x такой, что $\|x - \hat{x}\|_{C^1([t_0,\, t_1],\, \mathbb{R}^n)} < \varepsilon$, выполнено $\mathcal{L}_0(x) \ge \mathcal{L}_0(\hat{x})$.

Пусть для любого $j = \overline{0,\, m}$ и для любого $i = \overline{1,\, n}$ функции $(L_j)_{x_i} : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ и $(L_j)_{\dot{x}_i} : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ непрерывны. Тогда, как уже доказывалось, функционалы $\mathcal{L}_j : C^1([t_0,\, t_1],\, \mathbb{R}) \to \mathbb{R}$ непрерывно дифференцируемы.

Таким образом, задача (30) является примером гладкой задачи с ограничениями типа равенств. В качестве пространства X берется $C^1([t_0,\, t_1],\, \mathbb{R}^n)$, в качестве Y — пространство $\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$, $f_0(x) := \mathcal{L}_0(x)$, $F(x) = (\mathcal{L}_1(x) - c_1,\, \ldots,\, \mathcal{L}_m(x) - c_m,\, x(t_0) - x_{(0)},\, x(t_1) - x_{(1)})$. Так как Y конечномерно, то для любой точки $x \in X$ образ оператора $F'(x)$ будет замкнут.

Пусть $\hat{x}$ — точка локального минимума в (30). В силу принципа Лагранжа, существуют числа $\lambda_0,\, \ldots,\, \lambda_m,\, \mu_1,\, \ldots,\, \mu_n,\, \nu_1,\, \ldots,\, \nu_n$, одновременно не равные 0 и такие, что $\mathcal{L}'(\hat{x}) = 0$, где

$$\mathcal{L}(x) = \sum_{j=0}^{m} \lambda_j \mathcal{L}_j(x) + \sum_{i=1}^{n} (\mu_i x_i(t_0) + \nu_i x_i(t_1))$$

(константы c_j и $x_{(0)}$, $x_{(1)}$ можно не писать, так как их производная равна 0). Функционал $\mathcal{L}$ имеет такой же вид, как в задаче Больца, поэтому его производная равна 0 тогда и только тогда, когда выполнены уравнения Эйлера–Лагранжа и условия трансверсальности. При этом интегральная часть функционала $\mathcal{L}$ имеет вид $\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt$, где $L = \sum\limits_{j=0}^{m} \lambda_j L_j$.

Отметим также, что если $(\lambda_0, \ldots, \lambda_m) = (0, \ldots, 0)$, то из условий трансверсальности получим $\mu_i = 0$, $\nu_i = 0$, $1 \le i \le n$, то есть все множители Лагранжа окажутся равными 0. Значит, этот случай невозможен.

Тем самым, получается

Теорема 12. *Пусть $\hat{x}$ — точка локального минимума в задаче* (30). *Тогда существует набор чисел $(\lambda_0, \ldots, \lambda_m)$, одновременно не равных* 0, *таких, что для лагранжиана $L = \sum\limits_{j=0}^{m} \lambda_j L_j$ выполнено уравнение Эйлера–Лагранжа: $-\frac{d}{dt} L_{\dot{x}} + L_x = 0$.*

Замечание. Если функции L_j определены на $[t_0, t_1] \times U \times V$, где U, $V \subset \mathbb{R}^n$ — непустые открытые множества, и если в задаче (30) дополнительно написать ограничения $x(t) \in U$, $\dot{x}(t) \in V$, то необходимые условия локального минимума формулируются так же.

Задача 17. Пусть $l > 0$. Доказать, что допустимые экстремали в задаче

$$\begin{gathered} \int\limits_0^1 (y\dot{x} - x\dot{y})\, dt \to \max, \quad \int\limits_0^1 \sqrt{\dot{x}^2 + \dot{y}^2}\, dt = l, \\ x(0) = x(1) = y(0) = y(1) = 0, \quad \dot{x}^2 + \dot{y}^2 > 0 \end{gathered} \tag{31}$$

являются параметризацией окружности (пройденной один или несколько раз).

В этой задаче ищется замкнутая кривая заданной длины, проходящая через точку (0, 0), ограничивающая фигуру максимальной площади (задача Дидоны). Существование точки минимума мы пока не доказываем. Позже будет доказана теорема существования точки минимума, и (31) будет сведена к эквивалентной задаче, у которой существование минимума уже будет нетрудно доказать.

9 Задача выпуклого программирования

Пусть X — линейное пространство, $f_0, \ldots, f_m : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции. Задача

$$\left\{ \begin{array}{l} f_0(x) \to \inf, \\ f_j(x) \le 0, \quad 1 \le j \le m, \end{array} \right. \tag{32}$$

называется задачей выпуклого программирования.

Точка $x \in X$ называется допустимой, если $f_j(x) \le 0$, $1 \le j \le m$. Заметим, что множество допустимых точек выпукло.

Допустимая точка $\hat{x}$ называется точкой минимума в задаче (32), если $f_0(\hat{x}) \le f_0(x)$ для любого допустимого x.

Теорема Каруша – Куна – Таккера является аналогом принципа Лагранжа.

Теорема 13. (Каруш – Кун – Таккер). *Пусть X — линейное пространство, $f_0, \ldots, f_m : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции.*

1. *(необходимое условие). Пусть $\hat{x}$ — точка минимума в задаче* (32). *Тогда существует ненулевой набор чисел $\lambda_0, \lambda_1, \ldots, \lambda_m$ со следующими свойствами:*
   - *(a) $\lambda_j \ge 0$, $0 \le j \le m$ (условие неотрицательности);*
   - *(b) $\lambda_j f_j(\hat{x}) = 0$, $1 \le j \le m$ (условие дополняющей нежесткости);*
   - *(c) $\hat{x}$ является точкой минимума функции $\mathcal{L}(x) := \sum\limits_{j=0}^{m} \lambda_j f_j(x)$ (условие минимума).*
2. *(достаточное условие). Пусть $\hat{x}$ — допустимая точка. Пусть существует набор чисел $\lambda_0, \lambda_1, \ldots, \lambda_m$ со свойствами a)–c), при этом $\lambda_0 > 0$. Тогда $\hat{x}$ — точка минимума в задаче* (32).
3. *Пусть существует точка $\overline{x} \in X$ такая, что $f_j(\overline{x}) < 0$, $1 \le j \le m$ (условие Слейтера). Тогда, если $\lambda_0, \lambda_1, \ldots, \lambda_m$ — ненулевой набор чисел со свойствами a)–c), то $\lambda_0 > 0$.*

Таким образом, если выполнено условие Слейтера, то $\lambda_0 > 0$ и необходимое условие минимума совпадает с достаточным.

Задача 18. Привести пример задачи выпуклого программирования такой, что допустимая $\hat{x}$ — не точка минимума, но существует ненулевой набор $(\lambda_0, \ldots, \lambda_m)$, удовлетворяющий a)–c) из теоремы Куна – Таккера.

Теперь докажем теорему.

Доказательство. Необходимое условие. Пусть $\hat{x}$ — точка минимума. Без ограничения общности можно считать, что $f_0(\hat{x}) = 0$. Рассмотрим множества в $\mathbb{R}^{m+1}$:

$$A = \{(\mu_0, \mu_1, \ldots, \mu_m) : \exists x \in X : \ f_j(x) \le \mu_j, \ 0 \le j \le m\},$$

$$B = \{(\mu_0,\, \mu_1,\, \ldots,\, \mu_m) :\, \mu_j < 0,\; 0 \le j \le m\}.$$

Множество B выпукло и открыто. Докажем, что A выпукло. В самом деле, пусть $(\mu_0,\, \mu_1,\, \ldots,\, \mu_m)$, $(\nu_0,\, \nu_1,\, \ldots,\, \nu_m) \in A$, $\lambda \in [0,\, 1]$. По определению, существуют точки x, $y \in X$ такие, что $f_j(x) \le \mu_j$, $f_j(y) \le \nu_j$, $0 \le j \le m$. Так как функции f_j выпуклы, то $f_j((1-\lambda)x + \lambda y) \le (1-\lambda)f_j(x) + \lambda f_j(y) \le (1-\lambda)\mu_j + \lambda\nu_j$. Значит, $(1-\lambda)(\mu_0,\, \mu_1,\, \ldots,\, \mu_m) + \lambda(\nu_0,\, \nu_1,\, \ldots,\, \nu_m) \in A$.

Докажем, что множества A и B не пересекаются. В самом деле, пусть $(\mu_0,\, \mu_1,\, \ldots,\, \mu_m) \in A \cap B$. Тогда существует $x \in X$ такое, что $f_j(x) \le \mu_j < 0$, $0 \le j \le m$. Значит, точка x допустимая и $f_0(x) < 0$. Это противоречит тому, что минимум в задаче (32) равен 0.

По теореме отделимости, существует набор чисел $(\lambda_0,\, \lambda_1,\, \ldots,\, \lambda_m) \ne (0,\, 0,\, \ldots,\, 0)$ такой, что

$$\inf_{(\mu_0,\, \mu_1,\, \ldots,\, \mu_m) \in A} \sum_{j=0}^{m} \lambda_j \mu_j \ge \sup_{(\mu_0,\, \mu_1,\, \ldots,\, \mu_m) \in B} \sum_{j=0}^{m} \lambda_j \mu_j. \tag{33}$$

Пусть $\lambda_{j_*} < 0$ для некоторого $j_* \in \{0,\, \ldots,\, m\}$. Зафиксируем какие-нибудь отрицательные числа μ_j при $j \ne j_*$; μ_{j_*} устремим к $-\infty$ и получим, что правая часть (33) равна $+\infty$. Значит, $\lambda_j \ge 0$, $0 \le j \le m$ (т.е. выполнено условие неотрицательности). Тогда правая часть (33) равна 0, т.е.

$$\sum_{j=0}^{m} \lambda_j \mu_j \ge 0, \quad (\mu_0,\, \mu_1,\, \ldots,\, \mu_m) \in A. \tag{34}$$

В частности, так как $(f_0(x),\, f_1(x),\, \ldots,\, f_m(x)) \in A$ для любого $x \in X$, то

$$\sum_{j=0}^{m} \lambda_j f_j(x) \ge 0. \tag{35}$$

Проверим условие дополняющей нежесткости. Пусть $1 \le j \le m$. Положим $\mu_i = 0$, $i \ne j$, $\mu_j = f_j(\hat{x})$. Тогда $f_0(\hat{x}) = 0 = \mu_0$, $f_i(\hat{x}) \le 0 = \mu_i$, $i \in \{1,\, \ldots,\, m\} \backslash \{j\}$, $f_j(\hat{x}) = \mu_j$. Значит, $(\mu_0,\, \mu_1,\, \ldots,\, \mu_m) \in A$ и поэтому $\lambda_j f_j(\hat{x}) \overset{(34)}{\ge} 0$. С другой стороны, $\lambda_j \ge 0$, $f_j(\hat{x}) \le 0$. Значит, $\lambda_j f_j(\hat{x}) = 0$.

Из условия дополняющей нежесткости и равенства $f_0(\hat{x}) = 0$ получаем $\sum\limits_{j=0}^{m} \lambda_j f_j(\hat{x}) = 0$. Это вместе с (35) дает условие минимума.

Достаточное условие. Пусть $\hat{x}$ — допустимая точка, $(\lambda_0, \lambda_1, \ldots, \lambda_m)$ — набор чисел со свойствами a)–c), $\lambda_0 > 0$. Покажем, что $f_0(x) \geq f_0(\hat{x})$ для любой допустимой точки x. Из неравенства $\mathcal{L}(x) \geq \mathcal{L}(\hat{x})$ и условия дополняющей нежесткости получаем

$$\sum_{j=0}^{m} \lambda_j f_j(x) \geq \sum_{j=0}^{m} \lambda_j f_j(\hat{x}) = \lambda_0 f_0(\hat{x}).$$

Так как $\lambda_j \geq 0$, $f_j(x) \leq 0$ при $1 \leq j \leq m$, то $\lambda_0 f_0(x) \geq \lambda_0 f_0(\hat{x})$. Остается сократить на $\lambda_0 > 0$.

Условие строгой положительности λ_0. Пусть выполнено условие Слейтера, но $\lambda_0 = 0$. Из условий минимума и дополняющей нежесткости получаем

$$\sum_{j=1}^{m} \lambda_j f_j(\overline{x}) \geq \sum_{j=1}^{m} \lambda_j f_j(\hat{x}) = 0.$$

С другой стороны, левая часть строго отрицательна: $\lambda_j \geq 0$, $f_j(\overline{x}) < 0$ для любого $j = 1, \ldots, m$, при этом λ_j одновременно не равны нулю. Получили противоречие. $\square$

10 Гладкая задача с ограничениями типа равенств и неравенств

Пусть X, Y — нормированные пространства, $U \subset X$ — открытое множество, $f_0, \ldots, f_m : U \to \mathbb{R}$, $F : U \to Y$. Рассмотрим задачу

$$\begin{cases} f_0(x) \to \inf, \\ f_j(x) \leq 0, \quad 1 \leq j \leq m, \\ F(x) = 0, \\ x \in U. \end{cases} \tag{36}$$

Точка $x \in U$ называется допустимой для задачи (36), если $f_j(x) \leq 0$ для любого $j = 1, \ldots, m$ и $F(x) = 0$. Допустимая точка $\hat{x}$ называется точкой локального минимума в задаче (36), если существует $\varepsilon > 0$ такое, что для любой допустимой точки x такой, что $\|x - \hat{x}\| < \varepsilon$, выполнено $f_0(x) \geq f_0(\hat{x})$.

Теорема 14. *Пусть пространства X, Y полны, отображения $f_0, \ldots, f_m$, F строго дифференцируемы в точке $\hat{x}$. Пусть $\operatorname{Im} F'(\hat{x})$ замкнут. Тогда существуют числа $\lambda_0, \ldots, \lambda_m \in \mathbb{R}$ и функционал $y^* \in Y^*$, одновременно не равные нулю, удовлетворяющие следующим условиям:*

1. *$\lambda_j \geq 0$, $0 \leq j \leq m$ (условие неотрицательности);*
2. *$\lambda_j f_j(\hat{x}) = 0$, $1 \leq j \leq m$ (условие дополняющей нежесткости);*
3. *$\mathcal{L}'(\hat{x}) = 0$, где $\mathcal{L}(x) = \sum\limits_{j=0}^{m} \lambda_j f_j(x) + y^*(F(x))$ (условие стационарности).*

Условие стационарности означает, что для любого $h \in X$ выполнено

$$\sum_{j=0}^{m} \lambda_j f_j'(\hat{x})[h] + y^*(F'(\hat{x})[h]) = 0. \tag{37}$$

Доказательство. Сначала заметим, что достаточно рассмотреть случай, когда $f_j(\hat{x}) = 0$ для любого $j \in \{1, \ldots, m\}$. В самом деле, если для некоторого $j \in \{1, \ldots, m\}$ выполнено $f_j(\hat{x}) < 0$, то в силу непрерывности f_j в точке $\hat{x}$ получаем, что $f_j(x) < 0$ в некоторой окрестности $\hat{x}$. Поэтому вместо множества U можно рассмотреть его пересечение с этой окрестностью и получить задачу с меньшим числом ограничений. Если положить $\lambda_i = 0$ для всех i таких, что $f_i(\hat{x}) < 0$, то

$$\sum_{j=0}^{m} \lambda_j f_j(x) + y^*(F(x)) = \sum_{i:\, f_i(x)=0} \lambda_i f_i(x) + y^*(F(x)).$$

Также без ограничения общности можно считать, что $f_0(\hat{x}) = 0$.

Случай 1: $\operatorname{Im} F'(\hat{x}) = Y$. Рассмотрим множество

$$C = \{(\mu_0, \ldots, \mu_m, y) \in \mathbb{R}^{m+1} \times Y : \exists x \in X : \ f_j'(\hat{x})[x] < \mu_j, \ 0 \leq j \leq m, \ F'(\hat{x})[x] = y\}.$$

Множество C выпукло.

Покажем, что $0 \notin C$. В самом деле, иначе существует точка x_0 такая, что

$$f_j'(\hat{x})[x_0] < 0, \ 0 \leq j \leq m, \quad F'(\hat{x})[x_0] = 0. \tag{38}$$

Последнее соотношение означает, что $x_0 \in \ker F'(\hat{x})$. По теореме о касательном пространстве, $x_0 \in T_{\hat{x}}M$, где $M = \{x \in U : \ F(x) = 0\}$. Значит, существуют $\delta > 0$ и отображение $r : (-\delta, \delta) \to Y$ такие, что $r(t) = o(t)$ при $t \to 0$ и $\hat{x} + tx_0 + r(t) \in M$ для любого $t \in (-\delta, \delta)$, т.е. $F(\hat{x} + tx_0 + r(t)) = 0$. Далее,

$$f_j(\hat{x} + tx_0 + r(t)) = f_j(\hat{x}) + f_j'(\hat{x})[tx_0 + r(t)] + o(t) = tf_j'(\hat{x})[x_0] + o(t).$$

Значит, при малых $t > 0$ выполнено $f_j(\hat{x} + tx_0 + r(t)) \overset{(38)}{<} 0$, т.е. $\hat{x} + tx_0 + r(t)$ допустимая и $f_0(\hat{x} + tx_0 + r(t)) < 0$. Но тогда $\hat{x}$ не является точкой локального минимума.

Теперь покажем, что $\operatorname{int} C \neq \varnothing$. Для $d > 0$ рассмотрим множество

$$C_d = \{(\mu_0, \ldots, \mu_m, y) \in \mathbb{R}^{m+1} \times Y : \ \mu_j > d, \ 0 \le j \le m, \ y \in F'(\hat{x})(B_X)\}.$$

Так как $F'(\hat{x})$ сюръективно, то $F'(\hat{x})(B_X)$ содержит окрестность нуля (по теореме Банаха об открытом отображении). Если d достаточно велико, то $C_d \subset C$. В самом деле, для любого $x \in B_X$ выполнено $f_j'(\hat{x})[x] \le \|f_j'(\hat{x})\|$; значит, достаточно взять $d > \max_{0 \le j \le m} \|f_j'(\hat{x})\|$.

Таким образом, можно применить теорему отделимости. Получаем, что существуют $\lambda_0, \ldots, \lambda_m \in \mathbb{R}$, $y^* \in Y^*$, одновременно не равные 0, такие, что

$$\sum_{j=0}^m \lambda_j \mu_j + y^*(y) \ge 0, \quad (\mu_0, \ldots \mu_m, y) \in C.$$

Пусть $y = 0$, μ_j — произвольные положительные числа. Тогда $f_j'(\hat{x})[0] < \mu_j$, $0 \le j \le m$, $F'(\hat{x})[0] = y$. Значит, $(\mu_0, \ldots, \mu_m, 0) \in C$ и $\sum\limits_{j=0}^m \lambda_j \mu_j \ge 0$. Возьмем $\mu_k = 1$, $\mu_j = \varepsilon$, $j \neq k$. Устремив ε к 0, получим $\lambda_k \ge 0$. Таким образом, получили условие неотрицательности.

Теперь положим $\mu_j = f_j'(\hat{x})[h] + \varepsilon$, $0 \le j \le m$, $y = F'(\hat{x})[h]$, где $h \in X$, $\varepsilon > 0$. Тогда $(\mu_0, \ldots, \mu_m, y) \in C$ и

$$\sum_{j=0}^m \lambda_j (f_j'(\hat{x})[h] + \varepsilon) + y^*(F'(\hat{x})[h]) \ge 0.$$

Устремив $\varepsilon \to 0$, получаем

$$\sum_{j=0}^m \lambda_j f_j'(\hat{x})[h] + y^*(F'(\hat{x})[h]) \ge 0, \quad h \in X.$$

Заменив h на $-h$, получим противоположное неравенство. Значит,

$$\sum_{j=0}^{m} \lambda_j f_j'(\hat{x})[h] + y^*(F'(\hat{x})[h]) = 0, \quad h \in X.$$

Таким образом, получили (37).

Случай 2: $\operatorname{Im} F'(\hat{x}) \neq Y$. По лемме о нетривиальном аннуляторе, существует $y^* \in Y^* \backslash \{0\}$ такой, что $y^*(y) = 0$ для любого $y \in \operatorname{Im} F'(\hat{x})$. Взяв $\lambda_0 = \cdots = \lambda_m = 0$, получаем требуемый результат. □

11 Достаточное условие глобального минимума в задаче с ограничениями типа равенств и неравенств; достаточное условие единственности точки минимума

Пусть X, Y — линейные пространства, $U \subset X$, $f_0, \ldots, f_m : U \to \mathbb{R}$, $F : U \to Y$. Рассмотрим задачу

$$\begin{cases} f_0(x) \to \inf, \\ f_j(x) \le 0, \quad 1 \le j \le m, \\ F(x) = 0, \\ x \in U. \end{cases} \tag{39}$$

Следующее утверждение дает достаточное условие минимума в задаче (39).

Предложение 20. *Пусть $\hat{x}$ — допустимая точка. Предположим, что существуют числа $\lambda_0, \ldots, \lambda_m$ и линейный функционал y^* на пространстве Y со следующими свойствами:*

1. $\lambda_0 > 0$, $\lambda_1 \ge 0, \ldots, \lambda_m \ge 0$;
2. $\lambda_j f_j(\hat{x}) = 0$, $1 \le j \le m$;
3. $\mathcal{L}(\hat{x}) = \min_{x \in U} \mathcal{L}(x)$, *где* $\mathcal{L}(x) = \sum_{j=0}^{m} \lambda_j f_j(x) + y^*(F(x))$.

Тогда $\hat{x}$ — точка глобального минимума в (39).

Доказательство. Пусть x — допустимая точка. Из неравенства $\mathcal{L}(x) \geq \mathcal{L}(\hat{x})$ получаем

$$\sum_{j=0}^{m} \lambda_j f_j(x) + y^*(F(x)) \geq \sum_{j=0}^{m} \lambda_j f_j(\hat{x}) + y^*(F(\hat{x})).$$

Так как $F(x) = F(\hat{x}) = 0$ и $\lambda_j f_j(\hat{x}) = 0$, получаем

$$\sum_{j=0}^{m} \lambda_j f_j(x) \geq \lambda_0 f_0(\hat{x}).$$

Из неравенств $\lambda_j \geq 0$ и $f_j(\hat{x}) \leq 0$ получаем $\lambda_0 f_0(x) \geq \lambda_0 f_0(\hat{x})$. Так как $\lambda_0 > 0$, то $f_0(x) \geq f_0(\hat{x})$. $\square$

Определение 2. *Пусть X — линейное пространство, $U \subset X$ — выпуклое множество. Функция $f : U \to \mathbb{R}$ называется строго выпуклой, если для любых различных точек $x_1,\ x_2 \in U$ и любого $\lambda \in (0,\,1)$ выполнено $f((1-\lambda)x_1 + \lambda x_2) < (1-\lambda)f(x_1) + \lambda f(x_2)$.*

Предложение 21. *Пусть в задаче*

$$\begin{cases} f_0(x) \to \inf, \\ x \in U \end{cases}$$

множество U выпуклое, функция f_0 строго выпуклая. Тогда точка минимума единственна.

Доказательство. Пусть $x_1,\ x_2 \in U$ — две различные точки минимума. Тогда $\frac{x_1+x_2}{2} \in U$, при этом $f_0\left(\frac{x_1+x_2}{2}\right) < \frac{f_0(x_1)+f_0(x_2)}{2} = \min_{x\in U} f(x)$ — противоречие. $\square$

Задача 19. (распределение с максимальной энтропией). Пусть $\rho : [0,\,+\infty) \to (0,\,+\infty)$, $\int\limits_0^\infty \rho(x)\,dx = 1$ (функция ρ имеет смысл плотности распределения). Энтропией называется величина $S = -\int\limits_0^\infty \rho(x)\ln\rho(x)\,dx$. Найти функцию ρ, для которой энтропия максимальна при заданном среднем (т.е. задано ограничение $\int\limits_0^\infty x\rho(x)\,dx = C_1$).

12 Задача Лагранжа

12.1 Постановка задачи и формулировка основного результата

Пусть $n,\, s \in \mathbb{N}$, $m \in \mathbb{Z}_+$, $t_0 < t_1$, $L_j : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$ $(0 \le j \le m)$, $\varphi : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^n$ — непрерывные отображения. Кроме того, предполагаем, что для любого $i = 1,\, \dots,\, n$, $j = 1,\, \dots,\, s$, $k = 0,\, \dots,\, m$ существуют частные производные

$$(L_k)_{\xi_i}(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s), \quad (L_k)_{\eta_j}(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s),$$

$$\varphi_{\xi_i}(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s), \quad \varphi_{\eta_j}(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s),$$

при этом отображения

$$(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s) \mapsto (L_k)_{\xi_i}(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s),$$

$$(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s) \mapsto (L_k)_{\eta_j}(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s),$$

$$(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s) \mapsto \varphi_{\xi_i}(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s),$$

$$(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s) \mapsto \varphi_{\eta_j}(t,\, \xi_1,\, \dots,\, \xi_n,\, \eta_1, \dots,\, \eta_s)$$

непрерывны.

Пусть также для $0 \le k \le m$ заданы непрерывно-дифференцируемые отображения $l_k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.

Задача Лагранжа имеет вид:

$$\begin{cases} \int\limits_{t_0}^{t_1} L_0(t,\, x(t),\, u(t))\, dt + l_0(x(t_0),\, x(t_1)) \to \inf, \\ \int\limits_{t_0}^{t_1} L_k(t,\, x(t),\, u(t))\, dt + l_k(x(t_0),\, x(t_1)) \le 0, \quad 1 \le k \le m', \\ \int\limits_{t_0}^{t_1} L_k(t,\, x(t),\, u(t))\, dt + l_k(x(t_0),\, x(t_1)) = 0, \quad m'+1 \le k \le m, \\ \dot{x}(t) = \varphi(t,\, x(t),\, u(t)), \\ x \in C^1([t_0,\, t_1],\, \mathbb{R}^n),\ u \in C([t_0,\, t_1],\, \mathbb{R}^s). \end{cases} \tag{40}$$

Отметим, что функции $t \mapsto L_k(t,\, x(t),\, u(t))$ и $t \mapsto \varphi(t,\, x(t),\, u(t))$ непрерывны.

Обозначим $\mathcal{L}_k(x,\, u) = \int\limits_{t_0}^{t_1} L_k(t,\, x(t),\, u(t))\, dt + l_k(x(t_0),\, x(t_1))$, $0 \le k \le m$.

Пара $(x,\, u) \in C^1([t_0,\, t_1],\, \mathbb{R}^n) \times C([t_0,\, t_1],\, \mathbb{R}^s)$ называется допустимой, если $\mathcal{L}_k(x,\, u) \le 0$, $1 \le k \le m'$, $\mathcal{L}_k(x,\, u) = 0$, $m'+1 \le k \le m$, и $\dot{x}(t) = \varphi(t,\, x(t),\, u(t))$, $t \in [t_0,\, t_1]$.

Допустимая пара $(\hat{x},\, \hat{u})$ называется точкой локального минимума, если существует $\varepsilon > 0$ такое, что для любой допустимой пары $(x,\, u)$ такой, что $\|x - \hat{x}\|_{C^1([t_0,\, t_1],\, \mathbb{R}^n)} < \varepsilon$, $\|u - \hat{u}\|_{C([t_0,\, t_1],\, \mathbb{R}^s)} < \varepsilon$, выполнено $\mathcal{L}_0(\hat{x},\, \hat{u}) \le \mathcal{L}_0(x,\, u)$.

Перед тем, как сформулировать основной результат, введем обозначение. Для $p(\cdot) : [t_0,\, t_1] \to \mathbb{R}^n$, $p(t) = (p_1(t),\, \dots,\, p_n(t))$, положим

$$p(t)(\dot{x}(t) - \varphi(t,\, x(t),\, u(t))) = \sum_{j=1}^{n} p_j(t)(\dot{x}_j(t) - \varphi_j(t,\, x(t),\, u(t))).$$

Пусть $\lambda_0,\, \dots,\, \lambda_m \in \mathbb{R}$, $p(\cdot) \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$. Определим функцию Лагранжа

$$\mathcal{L}(x,\, u) = \sum_{k=0}^{m} \lambda_k \mathcal{L}_k(x,\, u) + \int_{t_0}^{t_1} p(t)(\dot{x}(t) - \varphi(t,\, x(t),\, u(t)))\, dt. \tag{41}$$

Выделим интегральное и внеинтегральное слагаемые:

$$\begin{gathered}\mathcal{L}(x,\, u) = \int_{t_0}^{t_1} L(t,\, x,\, \dot{x},\, u)\, dt + l(x(t_0),\, x(t_1)),\\ L(t,\, x,\, \dot{x},\, u) = \sum_{k=0}^{m} \lambda_k L_k(t,\, x,\, u) + p(t)(\dot{x} - \varphi(t,\, x,\, u)), \quad l = \sum_{k=0}^{m} \lambda_k l_k.\end{gathered} \tag{42}$$

Теорема 15. *Пусть $(\hat{x},\, \hat{u})$ — точка локального минимума в задаче* (40). *Тогда существует набор $(\lambda_0,\, \dots,\, \lambda_m,\, p(\cdot)) \neq (0,\, \dots,\, 0,\, 0)$, $\lambda_k \in \mathbb{R}$, $0 \le k \le m$, $p \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$, такой, что для функции $\mathcal{L}$, заданной равенствами* (41), (42), *выполнены следующие условия:*

1. *условие неотрицательности: $\lambda_j \ge 0$, $0 \le j \le m'$;*
2. *условие дополняющей нежесткости: $\lambda_j \mathcal{L}_j(\hat{x},\, \hat{u}) = 0$, $1 \le j \le m'$;*
3. *условия стационарности:*

(a) *по x, т.е. уравнение Эйлера–Лагранжа $-\frac{d}{dt}L_{\dot x}+L_x=0$ и условие трансверсальности*

$$L_{\dot x}(t_i,\,\hat x(t_i),\,\dot{\hat x}(t_i),\,\hat u(t_i))=(-1)^i\frac{\partial l}{\partial x(t_i)}(\hat x(t_0),\,\hat x(t_1));$$

(b) *по u: $L_u=0$.*

12.2 Задача Лагранжа как гладкая задача с ограничениями типа равенств

В качестве пространства X берется $C^1([t_0,\,t_1],\,\mathbb{R}^n)\times C([t_0,\,t_1],\,\mathbb{R}^s)$, в качестве Y — пространство $\mathbb{R}^{m-m'}\times C([t_0,\,t_1],\,\mathbb{R}^n)$. Ясно, что это банаховы пространства.

Проверим, что отображения $f_k(x,\,u):=\mathcal{L}_k(x,\,u)$, $0\le k\le m'$, и

$$F(x,\,u):=(\mathcal{L}_{m'+1}(x,\,u),\,\ldots,\,\mathcal{L}_m(x,\,u),\,\dot x(\cdot)-\varphi(\cdot,\,x(\cdot),\,u(\cdot)))$$

непрерывно дифференцируемы и вычислим их производные.

Отображение $(x,\,u)\mapsto\dot x$ является линейным непрерывным отображением из $C^1([t_0,\,t_1],\,\mathbb{R}^n)\times C([t_0,\,t_1],\,\mathbb{R}^s)$ в $C([t_0,\,t_1],\,\mathbb{R}^n)$. Значит, оно непрерывно дифференцируемо и его действие на вектор $(h,\,w)\in C^1([t_0,\,t_1],\,\mathbb{R}^n)\times C([t_0,\,t_1],\,\mathbb{R}^s)$ равно $\dot h$.

Отображение $(x,\,u)\mapsto\varphi(\cdot,\,x(\cdot),\,u(\cdot))$ является композицией отображения $(x,\,u)\mapsto(x,\,u)$ из $C^1([t_0,\,t_1],\,\mathbb{R}^n)\times C([t_0,\,t_1],\,\mathbb{R}^s)$ в $C([t_0,\,t_1],\,\mathbb{R}^n)\times C([t_0,\,t_1],\,\mathbb{R}^s)$ (это линейное непрерывное отображение) и оператора Немыцкого. Значит, оно дифференцируемо по Фреше и его производная имеет вид

$$(h,\,w)\mapsto\varphi_x(\cdot,\,x(\cdot),\,u(\cdot))h(\cdot)+\varphi_u(\cdot,\,x(\cdot),\,u(\cdot))w(\cdot);$$

из вида производной следует непрерывная дифференцируемость.

Аналогично вычисляется производная у отображения

$$(x,\,u)\mapsto L_k(\cdot,\,x(\cdot),\,u(\cdot));$$

взяв его композицию с интегралом по отрезку $[t_0,\,t_1]$ и продифференцировав отображение $(x,\,u)\mapsto l_k(x(t_0),\,x(t_1))$, получаем, что

$$\mathcal{L}'_k(x,\,u)[h,\,w]=\int\limits_{t_0}^{t_1}((L_k)_x(t,\,x(t),\,u(t))h(t)+(L_k)_u(t,\,x(t),\,u(t))w(t))\,dt+$$

$$+\frac{\partial l_k}{\partial x(t_0)}(x(t_0),\, x(t_1))h(t_0) + \frac{\partial l_k}{\partial x(t_1)}(x(t_0),\, x(t_1))h(t_1).$$

Отсюда нетрудно увидеть, что $(x,\, u) \mapsto \mathcal{L}_k(x,\, u)$ непрерывно дифференцируемо.

Таким образом, f_k и F непрерывно дифференцируемы.

Теперь покажем, что в каждой точке $(x,\, u)$ образ оператора $F'(x,\, u)$ замкнут.

Нам понадобится

Лемма 4. *Пусть X, Y — банаховы пространства, $A : X \to Y$, $B : X \to \mathbb{R}^n$ — линейные непрерывные операторы, A сюръективен. Тогда оператор $C : X \to Y \times \mathbb{R}^n$, $Cx = (Ax,\, Bx)$, имеет замкнутый образ.*

Доказательство. Пусть $\{x_n\}_{n\in\mathbb{N}} \subset X$, $(Ax_n,\, Bx_n) \underset{n\to\infty}{\to} (y,\, z)$. Покажем, что $(y,\, z) \in \operatorname{Im} C$.

Так как A сюръективен, то существует правый обратный оператор $R : Y \to X$ и число M такие, что $\|R\eta\| \le M\|\eta\|$, $\eta \in Y$. Положим $\xi_n = R(y - Ax_n)$. Тогда $\xi_n \underset{n\to\infty}{\to} 0$.

Пусть $\tilde{x}_n = x_n + \xi_n$. Тогда $A\tilde{x}_n = y$, $B\tilde{x}_n \underset{n\to\infty}{\to} z$.

Так как оператор A сюръективен, то $y = A\hat{x}$ для некоторого $\hat{x} \in X$. Значит, $\tilde{x}_n \in \hat{x} + \ker A$ для любого $n \in \mathbb{N}$. Множество $B(\hat{x} + \ker A)$ замкнуто (это сдвиг линейного подпространства в $\mathbb{R}^n$), то есть $z = B\xi$, $\xi \in \hat{x} + \ker A$. Тогда $A\xi = y$. Тем самым, $(A\xi,\, B\xi) = (y,\, z)$ и $(y,\, z) \in \operatorname{Im} C$. $\square$

Положим $A(h,\, w) = \dot{h} - \varphi_x(\cdot,\, x(\cdot),\, u(\cdot))h - \varphi_u(\cdot,\, x(\cdot),\, u(\cdot))w$. В силу леммы, нам достаточно показать, что оператор A сюръективен.

Пусть $y \in C([t_0,\, t_1],\, \mathbb{R}^n)$. Рассмотрим задачу

$$\dot{h}(t) - \varphi_x(t,\, x(t),\, u(t))h(t) = y(t), \quad h(t_0) = 0.$$

Это линейное неоднородное дифференциальное уравнение относительно h с непрерывными коэффициентами и непрерывной правой частью. Из курса дифференциальных уравнений известно, что решение этой задачи существует на всём отрезке $[t_0,\, t_1]$. Отсюда следует сюръективность оператора A.

Тем самым, выполнены все условия теоремы о необходимом условии локального минимума в гладкой задаче с ограничениями типа равенств и неравенств.

12.3 Заряды и сопряженное пространство к $C[t_0, t_1]$

Пусть Ω — множество, Σ — σ-алгебра измеримых подмножеств в Ω. Отображение $\mu : \Sigma \to \mathbb{R}$ называется зарядом, если для любых попарно не пересекающихся множеств $E_1, \ldots, E_n, \dots \in \Sigma$ выполнено

$$\mu\left(\sqcup_{n\in\mathbb{N}} E_n\right) = \sum_{n\in\mathbb{N}} \mu(E_n).$$

Напомним следующий факт из действительного анализа: существует разложение множества $\Omega = \Omega_+ \sqcup \Omega_-$ такое, что $\Omega_\pm \in \Sigma$ и для любого $E \in \Sigma$ выполнено $\mu(E) \ge 0$, если $E \subset \Omega_+$, и $\mu(E) \le 0$, если $E \subset \Omega_-$.

Пусть $E \in \Sigma$. Положим $\mu_\pm(E) = \pm\mu(E \cap \Omega_\pm)$. Тогда $\mu_\pm$ — σ-аддитивные меры, $\mu(E) = \mu_+(E) - \mu_-(E)$. Мера $|\mu| = \mu_+ + \mu_-$ называется вариацией заряда μ.

Пусть $f : \Omega \to \mathbb{R}$ — измеримая функция. Если $f|_{\Omega_\pm}$ интегрируемы по Лебегу относительно мер $\mu_\pm$, то f называется интегрируемой по Лебегу относительно заряда μ; интеграл Лебега определяется как

$$\int\limits_\Omega f\,d\mu = \int\limits_\Omega f\,d\mu_+ - \int\limits_\Omega f\,d\mu_-.$$

Отметим, что

$$\left|\int\limits_\Omega f\,d\mu\right| \le \int\limits_\Omega |f|\,d|\mu| < \infty.$$

Также выполнена аддитивность интеграла Лебега относительно зарядов: если μ_1, μ_2 — заряды, f интегрируема по Лебегу относительно μ_1 и μ_2, то f интегрируема по Лебегу относительно $\mu_1 + \mu_2$ и

$$\int\limits_\Omega f\,d(\mu_1+\mu_2) = \int\limits_\Omega f\,d\mu_1 + \int\limits_\Omega f\,d\mu_2.$$

Нам также понадобится утверждение о перестановке пределов интегрирования (оно выводится из теоремы Фубини для мер). Пусть Ω_j — множества, Σ_j — σ-алгебры на Ω_j, μ_j — заряды на (Ω_j, Σ_j), $j = 1, 2$. Пусть функция $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ интегрируема по Лебегу относительно $|\mu_1| \times |\mu_2|$. Тогда

$$\int\limits_{\Omega_1}\left(\int\limits_{\Omega_2} f\,d\mu_2\right) d\mu_1 = \int\limits_{\Omega_2}\left(\int\limits_{\Omega_1} f\,d\mu_1\right) d\mu_2.$$

Напомним, что борелевской σ-алгеброй на отрезке $[t_0,\, t_1]$ называется минимальная σ-алгебра, порожденная системой открытых множеств. Будем ее обозначать через $\mathcal{B}$.

Заметим, что любая непрерывная функция на $[t_0,\, t_1]$ интегрируема по Лебегу относительно заряда на $\mathcal{B}$ (т.к. она измерима и ограничена).

Сформулируем следующую теорему из функционального анализа (без доказательства).

Теорема 16. *Пусть $x^* \in (C[t_0,\, t_1])^*$. Тогда на $([t_0,\, t_1],\, \mathcal{B})$ существует единственный заряд μ такой, что для любой функции $x \in C[t_0,\, t_1]$ выполнено $x^*(x) = \int\limits_{[t_0,\, t_1]} x\, d\mu$.*

Напомним, что норма пространства $C([t_0,\, t_1],\, \mathbb{R}^n)$ имеет вид $\|x\|_C = \max\{\|x_1\|_C,\, \ldots,\, \|x_n\|_C\}$. Поэтому пространство $C([t_0,\, t_1],\, \mathbb{R}^n)$ можно рассмотреть как $(C[t_0,\, t_1])^n$. Значит, если $x^* \in (C([t_0,\, t_1],\, \mathbb{R}^n))^*$, то существует единственный набор зарядов $(\mu_1,\, \ldots,\, \mu_n)$ на $([t_0,\, t_1],\, \mathcal{B})$ такой, что

$$x^*(x_1,\, \ldots,\, x_n) = \sum_{j=1}^{n} \int\limits_{[t_0,\, t_1]} x_j\, d\mu_j.$$

12.4 Вывод необходимых условий локального минимума в задаче Лагранжа

Напомним, что функция Лагранжа в теореме выглядит следующим образом:

$$\tilde{\mathcal{L}}(x,\, u) = \int\limits_{t_0}^{t_1} \Big(f(t,\, x(t),\, u(t)) + \sum_{j=1}^{n} p_j(t)(\dot{x}_j(t) - \varphi_j(t,\, x(t),\, u(t))) \Big)\, dt +$$

$$+ l(x(t_0),\, x(t_1)),$$

где $f = \sum\limits_{j=0}^{m} \lambda_j L_j$, $l = \sum\limits_{j=0}^{m} \lambda_j l_j$. Значит, уравнения Эйлера имеют вид

$$-\dot{p}_k(t) + f_{x_k}(t,\, x,\, u) - \sum_{j=1}^{n} p_j(t)(\varphi_j)_{x_k}(t,\, x,\, u) = 0, \qquad (43)$$

условия трансверсальности —

$$p_k(t_0) = \frac{\partial l}{\partial x_k(t_0)}(x(t_0),\, x(t_1)), \quad p_k(t_1) = -\frac{\partial l}{\partial x_k(t_1)}(x(t_0),\, x(t_1)), \tag{44}$$

условие стационарности по u —

$$f_{u_k}(t,\, x,\, u) - \sum_{j=1}^{n} p_j(t)(\varphi_j)_{u_k}(t,\, x,\, u) = 0. \tag{45}$$

Если применить принцип Лагранжа для гладкой задачи с ограничениями типа равенств и неравенств, то получится следующее утверждение: существуют числа $\lambda_0,\, \ldots,\, \lambda_m \in \mathbb{R}$ и заряды $\mu_1,\, \ldots,\, \mu_n$, одновременно не равные 0, такие, что выполнены условия неотрицательности и дополняющей нежесткости, при этом $\mathcal{L}'(\hat{x},\, \hat{u})[h,\, w] = 0$ для любых $h \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$, $w \in C([t_0,\, t_1],\, \mathbb{R}^s)$, где

$$\mathcal{L}(x,\, u) = \int_{t_0}^{t_1} f(t,\, x(t),\, u(t))\, dt + \sum_{j=1}^{n} \int_{t_0}^{t_1} (\dot{x}_j(t) - \varphi_j(t,\, x(t),\, u(t)))\, d\mu_j + l(x(t_0),\, x(t_1)).$$

Здесь снова $f = \sum\limits_{j=0}^{m} \lambda_j L_j$, $l = \sum\limits_{j=0}^{m} \lambda_j l_j$.

Выпишем производную:

$$\mathcal{L}'(x,\, u)[h,\, w] = \int_{t_0}^{t_1} (f_x(t,\, x(t),\, u(t))h(t) + f_u(t,\, x(t),\, u(t))w(t))\, dt +$$

$$+\sum_{j=1}^{n} \int_{t_0}^{t_1} (\dot{h}_j(t) - (\varphi_j)_x(t,\, x(t),\, u(t))h(t) - (\varphi_j)_u(t,\, x(t),\, u(t))w(t))\, d\mu_j(t) +$$

$$+\frac{\partial l}{\partial x(t_0)}(x(t_0),\, x(t_1))h(t_0) + \frac{\partial l}{\partial x(t_1)}(x(t_0),\, x(t_1))h(t_1).$$

Положим $w = 0$, $h_j = 0$ при $j \neq k$, $h_k = z$, где $z \in C^1[t_0,\, t_1]$ — произвольная функция такая, что $z(t_0) = z(t_1) = 0$. Тогда получим

$$\int_{t_0}^{t_1} f_{x_k}(t,\, x,\, u)z(t)\, dt + \int_{t_0}^{t_1} \dot{z}(t)\, d\mu_k(t) - \sum_{j=1}^{n} \int_{t_0}^{t_1} (\varphi_j)_{x_k}(t,\, x,\, u)z(t)\, d\mu_j(t) = 0. \tag{46}$$

Наша ближайшая цель — привести это равенство к виду $\int_{t_0}^{t_1} \dot z\, d\tilde\mu = 0$ для некоторого заряда $\tilde\mu$ и применить обобщенную лемму Дюбуа-Реймона.

Положим

$$B(t) = \int_{t_0}^{t} f_{x_k}(\tau,\, x(\tau),\, u(\tau))\, d\tau.$$

Проинтегрировав по частям и учитывая, что $z(t_0) = z(t_1) = 0$, получаем

$$\int_{t_0}^{t_1} f_{x_k}(t,\, x,\, u) z(t)\, dt = -\int_{t_0}^{t_1} B(t)\dot z(t)\, dt.$$

Определим заряд μ равенством $\mu(E) = \sum_{j=1}^{n} \int_E (\varphi_j)_{x_k}(t,\, x(t),\, u(t))\, d\mu_j(t)$. Тогда

$$\sum_{j=1}^{n} \int_{t_0}^{t_1} (\varphi_j)_{x_k}(t,\, x,\, u) z(t)\, d\mu_j(t) = \int_{t_0}^{t_1} z(t)\, d\mu(t).$$

Положим $g(t) = \mu([t,\, t_1])$. Так как μ — разность неотрицательных мер, то g — разность монотонных функций. Применив формулу Ньютона–Лейбница и следствие из теоремы Фубини о перестановке интегралов, получаем

$$\int_{t_0}^{t_1} z(t)\, d\mu(t) = \int_{t_0}^{t_1} \left(\int_{t_0}^{t} \dot z(\tau)\, d\tau \right) d\mu(t) = \int_{t_0}^{t_1} \left(\int_{[\tau,\, t_1]} \dot z(\tau)\, d\mu(t) \right) d\tau = \int_{t_0}^{t_1} \dot z(\tau) g(\tau)\, d\tau.$$

В итоге (46) сводится к равенству

$$-\int_{t_0}^{t_1} B(t)\dot z(t)\, dt + \int_{t_0}^{t_1} \dot z(t)\, d\mu_k(t) - \int_{t_0}^{t_1} \dot z(t) g(t)\, dt = 0,$$

или $\int_{t_0}^{t_1} \dot z\, d\tilde\mu = 0$ с

$$\tilde\mu(E) = -\int_E B(t)\, dt + \mu_k(E) - \int_E g(t)\, dt.$$

В силу обобщенной леммы Дюбуа-Реймона, существует константа $c \in \mathbb{R}$ такая, что
$$\int\limits_{[t_0,\, t_1]} x(t)\, d\tilde{\mu} = c \int\limits_{[t_0,\, t_1]} x(t)\, dt, \quad x \in C[t_0,\, t_1];$$
так как заряд, порождающий линейный непрерывный функционал на $C[t_0,\, t_1]$, единственный, то
$$-\int\limits_{[t_0,\, t]} B(\tau)\, d\tau + \mu_k([t_0,\, t]) - \int\limits_{[t_0,\, t]} g(\tau)\, d\tau = c(t - t_0).$$
В левой части первое и третье слагаемое — абсолютно непрерывные функции, правая часть — тоже, поэтому $t \mapsto \mu_k([t_0,\, t])$ также абсолютно непрерывна. Обозначим через p_k ее производную п.в. Продифференцируем по t:
$$-B(t) + p_k(t) - g(t) = c,$$
или
$$-\int\limits_{t_0}^{t} f_{x_k}(\tau,\, x(\tau),\, u(\tau))\, d\tau + p_k(t) - \sum_{j=1}^{n} \int\limits_{[t,\, t_1]} (\varphi_j)_{x_k}(\tau,\, x(\tau),\, u(\tau)) p_j(\tau)\, d\tau = c.$$
Снова первое и третье слагаемое в левой части и правая часть абсолютно непрерывны, так что p_k абсолютно непрерывна и
$$-f_{x_k}(t,\, x(t),\, u(t)) + \dot{p}_k(t) + \sum_{j=1}^{n} (\varphi_j)_{x_k}(t,\, x(t),\, u(t)) p_j(t) = 0;$$
остается заметить, что это равенство совпадает с (43) и что $\dot{p}_k$ непрерывна.

Пусть теперь z на концах не обязательно зануляется. Учитывая, что μ_k имеет плотность p_k, получаем
$$\int\limits_{t_0}^{t_1} f_{x_k}(t,\, x,\, u) z(t)\, dt + \int\limits_{t_0}^{t_1} \dot{z}(t) p_k(t)\, dt - \sum_{j=1}^{n} \int\limits_{t_0}^{t_1} (\varphi_j)_{x_k}(t,\, x,\, u) z(t) p_j(t)\, dt +$$
$$+ \frac{\partial l}{\partial x_k(t_0)}(x(t_0),\, x(t_1)) z(t_0) + \frac{\partial l}{\partial x_k(t_1)}(x(t_0),\, x(t_1)) z(t_1) = 0.$$

Проинтегрировав по частям $\int\limits_{t_0}^{t_1} \dot z(t)p_k(t)\,dt$ и учитывая доказанное уже уравнение Эйлера (43), получаем

$$z(t_1)p_k(t_1)-z(t_0)p_k(t_0)+\frac{\partial l}{\partial x_k(t_0)}(x(t_0),\, x(t_1))z(t_0)+\frac{\partial l}{\partial x_k(t_1)}(x(t_0),\, x(t_1))z(t_1)=0.$$

Отсюда следует (44).

Теперь возьмем $h=0$, $w_j=0$ при $j\ne k$, $w_k=v$, где $v\in C[t_0,\, t_1]$ — произвольная непрерывная функция. Получим

$$\int\limits_{t_0}^{t_1} f_{u_k}(t,\, x,\, u)v(t)\,dt-\int\limits_{t_0}^{t_1}\sum_{j=1}^{n}(\varphi_j)_{u_k}(t,\, x,\, u)p_j(t)v(t)\,dt=0.$$

Отсюда следует (45).

13 Задача оптимального управления

13.1 Постановка задачи и формулировка необходимого условия сильного локального минимума

Пусть $t_0<t_1$. Введем пространства $PC([t_0,\, t_1],\, \mathbb{R}^n)$ (пространство кусочно непрерывных функций) и $PC^1([t_0,\, t_1],\, \mathbb{R}^n)$ (пространство кусочно гладких функций).

Пусть $x:[t_0,\, t_1]\to\mathbb{R}^n$.

1. Скажем, что $x\in PC([t_0,\, t_1],\, \mathbb{R}^n)$, если существуют точки $t_0=\tau_0<\tau_1<\dots<\tau_{m-1}<\tau_m=t_1$ такие, что функция $x|_{(\tau_j,\,\tau_{j+1})}$ непрерывна и существуют односторонние пределы $\lim\limits_{t\to\tau_j+}x(t)$ и $\lim\limits_{t\to\tau_{j+1}-}x(t)$ для любого $j=0,\, \dots,\, m-1$.

2. Скажем, что $x\in PC^1([t_0,\, t_1],\, \mathbb{R}^n)$, если существуют точки $t_0=\tau_0<\tau_1<\dots<\tau_{m-1}<\tau_m=t_1$ такие, что $x|_{[\tau_j,\,\tau_{j+1}]}\in C^1([\tau_j,\, \tau_{j+1}],\, \mathbb{R}^n)$ для любого $j=0,\, \dots,\, m-1$.

Пусть $n,\, s\in\mathbb{N}$, $U\subset\mathbb{R}^s$ — произвольное множество, $L_j:[t_0,\, t_1]\times\mathbb{R}^n\times\mathbb{R}^s\to\mathbb{R}$ $(0\le j\le m)$, $\varphi:[t_0,\, t_1]\times\mathbb{R}^n\times\mathbb{R}^s\to\mathbb{R}^n$ — непрерывные функции,

$l_j : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ — непрерывно дифференцируемое отображение. Пусть также существуют частные производные отображений

$$(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_s) \mapsto L_j(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_s), \quad 0 \le j \le m,$$

и

$$(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_s) \mapsto \varphi(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_s)$$

по ξ_i для любого $i = 1,\, \ldots,\, n$, при этом отображения

$$(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_s) \mapsto (L_j)_{\xi_i}(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_s)$$

и

$$(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_s) \mapsto \varphi_{\xi_i}(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_s)$$

непрерывны.

Рассмотрим задачу

$$\begin{cases} \int\limits_{t_0}^{t_1} L_0(t,\, x(t),\, u(t))\, dt + l_0(x(t_0),\, x(t_1)) \to \inf, \\ \int\limits_{t_0}^{t_1} L_j(t,\, x(t),\, u(t))\, dt + l_j(x(t_0),\, x(t_1)) \le 0, \quad 1 \le j \le m', \\ \int\limits_{t_0}^{t_1} L_j(t,\, x(t),\, u(t))\, dt + l_j(x(t_0),\, x(t_1)) = 0, \quad m'+1 \le j \le m, \\ \dot{x}(t) = \varphi(t,\, x(t),\, u(t)), \\ u(t) \in U, \quad t \in [t_0,\, t_1], \\ x \in PC^1([t_0,\, t_1],\, \mathbb{R}^n),\ u \in PC([t_0,\, t_1],\, \mathbb{R}^s). \end{cases} \tag{47}$$

Обозначим $\mathcal{L}_j(x,\, u) = \int\limits_{t_0}^{t_1} L_j(t,\, x(t),\, u(t))\, dt + l_j(x(t_0),\, x(t_1))$.

Точка $(x,\, u) \in PC^1([t_0,\, t_1],\, \mathbb{R}^n) \times PC([t_0,\, t_1],\, \mathbb{R}^s)$ называется допустимой, если $\mathcal{L}_j(x,\, u) \le 0$, $1 \le j \le m'$, $\mathcal{L}_j(x,\, u) = 0$, $m'+1 \le j \le m$, $\dot{x}(t) = \varphi(t,\, x(t),\, u(t))$, $u(t) \in U$ для любого $t \in [t_0,\, t_1]$.

Отличия от задачи Лагранжа здесь следующие: 1) функции L_j и φ могут быть не дифференцируемыми по u, 2) есть включение $u(t) \in U$ (при этом множество U произвольное), 3) функции x, u принадлежат более широким пространствам.

Локальный минимум также будет пониматься в другом смысле.

Определение 3. *Допустимая точка* $(\hat{x},\, \hat{u})$ *называется точкой сильного локального минимума в задаче* (47), *если существует* $\varepsilon > 0$ *такое, что для любых допустимых* $(x,\, u)$ *таких, что* $\|x - \hat{x}\|_C < \varepsilon$, *выполнено* $\mathcal{L}_0(x,\, u) \ge \mathcal{L}_0(\hat{x},\, \hat{u})$.

Для $\lambda_0,\, \dots,\, \lambda_m \in \mathbb{R}$, $p \in PC^1([t_0,\, t_1],\, \mathbb{R}^n)$ определим функцию Лагранжа

$$\begin{aligned}\mathcal{L}(x,\, u) = \sum_{j=0}^{m} \lambda_j \mathcal{L}_j(x,\, u) + \int\limits_{t_0}^{t_1} p(t)(\dot{x}(t) - \varphi(t,\, x(t),\, u(t)))\, dt = \\ = \int\limits_{t_0}^{t_1} L(t,\, x,\, \dot{x},\, u)\, dt + l(x(t_0),\, x(t_1)),\end{aligned} \tag{48}$$

где $L(t,\, x,\, \dot{x},\, u) = \sum\limits_{j=0}^{m} \lambda_j L_j(t,\, x,\, u) + p(t)(\dot{x} - \varphi(t,\, x,\, u))$, $l_j = \sum\limits_{j=0}^{m} \lambda_j l_j$ (т.е. функция $\mathcal{L}$ определяется так же, как в задаче Лагранжа; включение $u(t) \in U$ в ней не участвует).

Теорема 17. *Пусть* $(\hat{x},\, \hat{u})$ — *точка сильного локального минимума в задаче* (47). *Тогда существуют* $\lambda_0,\, \dots,\, \lambda_m \in \mathbb{R}$, $p \in PC^1([t_0,\, t_1],\, \mathbb{R}^n)$, *одновременно не равные* 0, *такие, что выполнены следующие условия:*

1. *условие неотрицательности:* $\lambda_j \ge 0$, $0 \le j \le m'$;
2. *условие дополняющей нежесткости:* $\lambda_j \mathcal{L}_j(\hat{x},\, \hat{u}) = 0$, $1 \le j \le m'$;
3. *условия минимума:*
   (a) *стационарность по* x: *уравнение Эйлера–Лагранжа* $-\frac{d}{dt} L_{\dot{x}} + L_x = 0$ *и условие трансверсальности* $L_{\dot{x}}(t_i,\, \hat{x}(t_i),\, \dot{\hat{x}}(t_i),\, u(t_i)) = (-1)^i \frac{\partial l}{\partial x(t_i)}(\hat{x}(t_0),\, \hat{x}(t_1))$, $i = 0,\, 1$.
   (b) *по* u: *для любого* $t \in (t_0,\, t_1)$ *такого, что* $\hat{u}$ *непрерывна в* t, *выполнено*

   $$\min_{v \in U} L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t),\, v) = L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t),\, \hat{u}(t))$$

   (последнее условие называется принципом максимума Понтрягина).

Задача 20. (аэродинамическая задача Ньютона). Найти допустимые экстремали в задаче

$$\begin{cases} \int\limits_0^{T_0} \frac{t}{1+u^2}\, dt \to \inf, \\ x(0) = 0,\ x(T_0) = \xi, \\ \dot{x} = u, \\ u \geq 0; \end{cases} \tag{49}$$

здесь $T_0 > 0$, $\xi > 0$ — заданные параметры. (Ответ для $\hat{x}(t)$ записывается в параметрическом виде: $x = x(v)$, $t = t(v)$.) Доказать, что допустимая экстремаль существует и единственна, и что она будет точкой глобального минимума в задаче (49).

Физический смысл этой задачи следующий: требуется найти форму тела вращения с заданной высотой и шириной, образующая которого является монотонно возрастающей функцией, чтобы при движении в воздухе сопротивление было минимальным.

Мы докажем теорему 17 в частном случае, для задачи со свободным концом, $n = 1$. Задача ставится следующим образом:

$$\begin{cases} \int\limits_{t_0}^{t_1} f(t,\, x,\, u)\, dt \to \inf, \\ x(t_0) = x_0, \\ \dot{x} = \varphi(t,\, x,\, u), \\ u(t) \in U, \\ x \in PC^1[t_0,\, t_1],\ u \in PC[t_0,\, t_1]. \end{cases} \tag{50}$$

Напишем функцию $\mathcal{L}$ с $\lambda_0 = 1$:

$$\mathcal{L}(x,\, u) = \int\limits_{t_0}^{t_1} (f(t,\, x,\, u) + p(t)(\dot{x} - \varphi(t,\, x,\, u)))\, dt + \lambda_1 x(t_0).$$

Уравнение Эйлера имеет вид

$$\dot{p} = -p(t)\varphi_x(t,\, \hat{x}(t),\, \hat{u}(t)) + f_x(t,\, \hat{x}(t),\, \hat{u}(t)), \tag{51}$$

условие трансверсальности —

$$p(t_0) = \lambda_1,\ p(t_1) = 0, \tag{52}$$

принцип максимума Понтрягина —

$$\min_{v\in U}(f(t,\,\hat{x}(t),\,v)-p(t)\varphi(t,\,\hat{x}(t),\,v))=f(t,\,\hat{x}(t),\,\hat{u}(t))-p(t)\varphi(t,\,\hat{x}(t),\,\hat{u}(t)),$$

т.е.

$$\begin{gathered} f(t,\,\hat{x}(t),\,v)-f(t,\,\hat{x}(t),\,\hat{u}(t))- \\ -p(t)(\varphi(t,\,\hat{x}(t),\,v)-\varphi(t,\,\hat{x}(t),\,\hat{u}(t)))\ge 0, \quad \forall v\in U. \end{gathered} \tag{53}$$

Перед тем, как доказать это утверждение, напомним подробнее некоторые факты из теории обыкновенных дифференциальных уравнений.

13.2 Теоремы существования, единственности и гладкости по начальному условию решений обыкновенных дифференциальных уравнений

Сначала напомним утверждение, которое доказывалось в курсе дифференциальных уравнений с применением теоремы о неподвижной точке для λ-сжимающего отображения.

Предложение 22. *Пусть $g : [0,\,T]\times\mathbb{R}\to\mathbb{R}$ — непрерывное отображение, $x_0\in\mathbb{R}$. Пусть существуют $R,\,M>0$, такие что для любого $t\in[0,\,T]$ функция $x\mapsto g(t,\,x)$ липшицева на $[x_0-R,\,x_0+R]$ с константой M. Тогда существует $\delta=\delta(R,\,M)>0$ такое, что для любого $\xi\in[x_0-R/2,\,x_0+R/2]$ решение $x(\cdot,\,\xi)$ задачи*

$$\dot{x}=g(t,\,x), \quad x(0)=\xi \tag{54}$$

существует на отрезке $[0,\,\delta]$. При этом

$$\|x(\cdot,\,\xi)-x(\cdot,\,\eta)\|_{C[0,\,\delta]}\le C|\xi-\eta|, \quad \xi,\,\eta\in[x_0-R/2,\,x_0+R/2],$$

где $C=C(R,\,M)$. Кроме того, существует $\tilde{C}(x_0,\,R,\,M)>0$ такое, что $\|x(\cdot,\,\xi)\|_{C[0,\,\delta]}\le\tilde{C}(x_0,\,R,\,M)$.

Предложение 23. *Пусть функция $g(t,\,\cdot)$ липшицева с константой $M(R)$ на $[x_0-R,\,x_0+R]$ для каждого $R>0$, $t\in[0,\,T]$. Предположим, что решение задачи (54) с $\xi=x_0$ продолжается на весь отрезок $[0,\,T]$, $R_0=\|x(\cdot,\,x_0)-x_0\|_{C[0,\,T]}$. Тогда существует окрестность V точки x_0*

такая, что для любого $\xi \in V$ решение задачи (54) *продолжается на отрезок* $[0, T]$*; при этом*

$$\|x(\cdot, \xi) - x(\cdot, \eta)\|_{C[0,T]} \le C|\xi - \eta|, \quad \xi, \eta \in V, \tag{55}$$

где константа C зависит только от R_0 и $M(\cdot)$.

Доказательство. Пусть $\delta = \delta(3R_0, M(3R_0))$ такое, как в предложении 22. Решение $x(\cdot, \xi)$ существует на $[0, \delta]$ при $|\xi - x_0| < \varepsilon < R_0/2$, при этом оно липшицево по ξ в $C[0, \delta]$ с константой, зависящей от R_0, $M(\cdot)$. Если $\varepsilon > 0$ достаточно мало, то $|x(\delta, \xi) - x_0| < 3R_0/2$, поэтому решение продолжается на $[\delta, 2\delta]$, при этом оно снова липшицево по начальному условию с константой, зависящей только от R_0, $M(\cdot)$. Дальше аналогично продолжаем решение на $[2\delta, 3\delta]$, $[3\delta, 4\delta]$ и т.д., пока не дойдем до правого конца отрезка.

Неравенство (55) также следует из предложения 22. $\square$

При дополнительных условиях на функцию g есть дифференцируемость решения по начальному условию. (В курсе дифференциальных уравнений это доказывалось, но, возможно, при большей гладкости g, чем нам требуется.)

Предложение 24. *Пусть в каждой точке (t, x) существует производная $g_x(t, x)$, при этом отображение $(t, x) \mapsto g_x(t, x)$ непрерывно. Предположим, что решение задачи* (54) *с $\xi = x_0$ продолжается на весь отрезок $[0, T]$ (обозначим это решение через $\hat{x}$). Тогда существует окрестность V точки x_0 такая, что для любого $\xi \in V$ решение задачи* (54) *продолжается на отрезок $[0, T]$. Пусть отображение $F : V \to C[0, T]$ задано формулой $F(\xi) = x(\cdot, \xi)$. Тогда $F'(x_0) = y$, где y — решение задачи*

$$\dot{y} = g_x(t, \hat{x}(t))y, \quad y(0) = 1. \tag{56}$$

Доказательство. Для любого $R > 0$ функция $g(t, \cdot)$ липшицева на $[x_0 - R, x_0 + R]$ с константой $\|g_x\|_{C([0,T]\times[x_0-R, x_0+R])}$. Поэтому в силу предыдущего утверждения решение (54) продолжается на весь отрезок $[0, T]$ для всех ξ из некоторой окрестности точки x_0. Кроме того, отображение F липшицево в этой окрестности с некоторой константой $C > 0$.

Теперь покажем, что существует производная $F'(x_0)$, при этом она совпадает с y. Для этого достаточно показать, что

$$\|x(\cdot, x_0 + \lambda) - \hat{x}(\cdot) - \lambda y(\cdot)\|_{C[0,T]} = o(\lambda), \quad \lambda \to 0. \tag{57}$$

Пусть $G : C^1[0, T] \to C[0, T]$, $G(x)(t) = \dot{x}(t) - g(t, x(t))$. Мы уже доказывали, что это отображение непрерывно дифференцируемо, $G'(x)[h](t) = \dot{h}(t) - g_x(t, x(t))h(t)$, $G'(x)$ сюръективно. Пусть

$$N = \{x \in C^1[0, T] :\ G(x) = 0\}.$$

Условие $x \in N$ эквивалентно тому, что x — решение дифференциального уравнения $\dot{x} = g(t, x)$ (на всём отрезке $[0, T]$). По теореме о касательном пространстве, $T_{\hat{x}}N = \ker G'(\hat{x})$. Значит, $y \in T_{\hat{x}}N$. По определению, это означает, что существует функция $r : (-\varepsilon, \varepsilon) \to C^1[0, T]$ такая, что $\hat{x} + \lambda y + r(\lambda) \in N$ для любого $\lambda \in (-\varepsilon, \varepsilon)$ и $r(\lambda) = \underset{\lambda\to 0}{o}(\lambda)$. Пусть $\omega(\lambda) = r(\lambda)(0)$. Тогда $\hat{x} + \lambda y + r(\lambda) = x(\cdot, x_0 + \lambda + \omega(\lambda))$. Получаем:

$$\|x(\cdot, x_0+\lambda) - \hat{x} - \lambda y\|_{C[0,T]} \leq \|x(\cdot, x_0+\lambda) - \hat{x} - \lambda y - r(\lambda)\|_{C[0,T]} + \|r(\lambda)\|_{C[0,T]} =$$

$$= \|x(\cdot, x_0 + \lambda) - x(\cdot, x_0 + \lambda + \omega(\lambda))\|_{C[0,T]} + \|r(\lambda)\|_{C[0,T]} \leq$$

$$\leq C|\omega(\lambda)| + \|r(\lambda)\|_{C[0,T]} = \underset{\lambda\to 0}{o}(\lambda).$$

Предложение доказано. □

13.3 Доказательство принципа максимума Понтрягина для задачи со свободным концом

Пусть $(\hat{x}, \hat{u})$ — точка сильного минимума в задаче (50), $\tau \in (t_0, t_1)$ — точка непрерывности $\hat{u}$. Возьмем произвольное $v \in U$ и для достаточно малых $\lambda > 0$ положим

$$u_\lambda(t) = \begin{cases} \hat{u}(t), & t \in [t_0, \tau - \lambda] \cup [\tau, t_1], \\ v, & t \in (\tau - \lambda, \tau). \end{cases}$$

Через x_λ обозначим решение задачи

$$\dot{x}(t) = \varphi(t, x(t), u_\lambda(t)), \quad x(0) = x_0. \tag{58}$$

Пара (x_λ, u_λ) называется игольчатой вариацией. Заметим, что она является допустимой.

Обозначим $g_\lambda(t, x) = \varphi(t, x, u_\lambda(t))$. Существует разбиение отрезка $[t_0, t_1] = \{\Delta_i\}_{i=1}^m$ такое, что $g_\lambda|_{\Delta_i\times\mathbb{R}}$ непрерывна и $(g_\lambda)'_x|_{\Delta_i\times\mathbb{R}}$ непрерывна (функцию u_λ заменяем в концах отрезка Δ_i ее односторонним пределом);

при этом для любого $R > 0$ существует $M(R) > 0$ такое, что для любого $t \in [t_0,\, t_1]$, $\lambda \in [0,\, \tau - t_0]$ функция $g_\lambda(t,\, \cdot)$ липшицева с константой $M(R)$ на $[x_0 - R,\, x_0 + R]$.

Применим результаты предыдущего параграфа.

Лемма 5. *Существует* $\lambda_0 > 0$ *такое, что для любого* $\lambda \in (0,\, \lambda_0)$ *решение задачи* (58) *существует на всём отрезке* $[t_0,\, t_1]$*; при этом*

$$\|x_\lambda - \hat{x}\|_{C[t_0,\, t_1]} \underset{\lambda \to 0+}{\to} 0. \tag{59}$$

Доказательство. При $t \in [t_0,\, \tau - \lambda]$ значения $x_\lambda(t)$ и $\hat{x}(t)$ совпадают.

Решение задачи $\dot{x} = g_\lambda(t,\, x)$, $x(\tau - \lambda) = \hat{x}(\tau - \lambda)$ существует на интервале положительной длины, зависящей только от $\|\hat{x}\|_{C[t_0,\, t_1]}$ и $M(\cdot)$. Значит, при малых $\lambda > 0$ решение существует на отрезке $[\tau - \lambda,\, \tau]$. При этом $\|x_\lambda\|_{C[\tau-\lambda,\, \tau]}$ оценивается сверху константой, зависящей только от $\|\hat{x}\|_{C[t_0,\, t_1]}$ и $M(\cdot)$. Но тогда

$$\|x_\lambda(\cdot) - \hat{x}(\tau - \lambda)\|_{C[\tau-\lambda,\, \tau]} \le \lambda \max_{t \in [\tau - \lambda,\, \tau]} |g_\lambda(t,\, x_\lambda(t))| \le C_1 \lambda,$$

где C_1 зависит только от φ, v, $\|\hat{x}\|_{C[t_0,\, t_1]}$ и $M(\cdot)$. Отсюда получаем, что

$$\|x_\lambda - \hat{x}\|_{C[t_0,\, \tau]} \underset{\lambda \to 0+}{\to} 0.$$

На отрезке $[\tau,\, t_1]$ при малых λ решение существует в силу предложения 23 (последовательно продолжаем на каждый отрезок, где $\hat{u}$ непрерывна). При этом $\|x_\lambda - \hat{x}\|_{C[\tau,\, t_1]} \underset{\lambda \to 0+}{\to} 0$ в силу непрерывной зависимости от начального условия. $\square$

Положим $\xi_\lambda = x_\lambda(\tau)$. Тогда в силу формулы Ньютона–Лейбница

$$\left.\frac{d\xi_\lambda}{d\lambda}\right|_{\lambda=0} = \lim_{\lambda \to 0+} \frac{1}{\lambda} \int\limits_{\tau-\lambda}^{\tau} (\dot{x}_\lambda(t) - \dot{\hat{x}}(t))\, dt =$$

$$= \lim_{\lambda \to 0+} \frac{1}{\lambda} \int\limits_{\tau-\lambda}^{\tau} (\varphi(t,\, x_\lambda(t),\, v) - \varphi(t,\, \hat{x}(t),\, \hat{u}(t)))\, dt =$$

$$= \lim_{\lambda \to 0+} \frac{1}{\lambda} \int\limits_{\tau-\lambda}^{\tau} (\varphi(t,\, \hat{x}(t),\, v) - \varphi(t,\, \hat{x}(t),\, \hat{u}(t)))\, dt +$$

$$+\lim_{\lambda\to 0+}\frac{1}{\lambda}\int\limits_{\tau-\lambda}^{\tau}(\varphi(t,\,x_\lambda(t),\,v)-\varphi(t,\,\hat{x}(t),\,v))\,dt=$$

$$=\varphi(\tau,\,\hat{x}(\tau),\,v)-\varphi(\tau,\,\hat{x}(\tau),\,\hat{u}(\tau)).$$

Итак,

$$\left.\frac{d\xi_\lambda}{d\lambda}\right|_{\lambda=0}=\varphi(\tau,\,\hat{x}(\tau),\,v)-\varphi(\tau,\,\hat{x}(\tau),\,\hat{u}(\tau)). \tag{60}$$

Теперь вычислим производную $\left.\frac{dx_\lambda(t)}{d\lambda}\right|_{\lambda=0}$ при $t\in[\tau,\,t_1]$. Напомним, что если $x(\cdot,\,\xi)$ — решение задачи $\dot{x}=\varphi(t,\,x,\,\hat{u}(t))$, $x(\tau)=\xi$, то, в силу предложения 24, $x'_\xi(\cdot,\,\hat{x}(\tau))$ — решение задачи

$$\dot{h}=\varphi_x(t,\,\hat{x}(t),\,\hat{u}(t))h,\quad h(\tau)=1. \tag{61}$$

Так как x_λ — решение задачи

$$\dot{x}=\varphi(t,\,x,\,\hat{u}(t)),\quad x(\tau)=\xi_\lambda,$$

то

$$\left.\frac{dx_\lambda(t)}{d\lambda}\right|_{\lambda=0}=x'_\xi(\cdot,\,\hat{x}(\tau))\left.\frac{d\xi_\lambda}{d\lambda}\right|_{\lambda=0}=y(t) \tag{62}$$

— решение задачи

$$\dot{y}=\varphi_x(t,\,\hat{x}(t),\,\hat{u}(t))y,\quad y(\tau)\overset{(60),(61)}{=}\varphi(\tau,\,\hat{x}(\tau),\,v)-\varphi(\tau,\,\hat{x}(\tau),\,\hat{u}(\tau)) \tag{63}$$

(здесь использовалось, что дифференциальное уравнение в (61), (63) линейное).

Теперь перейдем непосредственно к доказательству необходимых условий сильного минимума в (50).

При малых $\lambda>0$ выполнено $\int\limits_{t_0}^{t_1}f(t,\,x_\lambda(t),\,u_\lambda(t))\,dt-\int\limits_{t_0}^{t_1}f(t,\,\hat{x}(t),\,\hat{u}(t))\,dt\geq 0$ в силу (59) и определения сильного минимума. Значит,

$$0\leq\lim_{\lambda\to 0+}\frac{1}{\lambda}\int\limits_{t_0}^{t_1}(f(t,\,x_\lambda(t),\,u_\lambda(t))-f(t,\,\hat{x}(t),\,\hat{u}(t)))\,dt=$$

$$= \lim_{\lambda\to 0+} \frac{1}{\lambda} \int\limits_{\tau-\lambda}^{\tau} (f(t,\, x_\lambda(t),\, v) - f(t,\, \hat{x}(t),\, \hat{u}(t)))\, dt +$$

$$+ \lim_{\lambda\to 0+} \frac{1}{\lambda} \int\limits_{\tau}^{t_1} (f(t,\, x_\lambda(t),\, \hat{u}(t)) - f(t,\, \hat{x}(t),\, \hat{u}(t)))\, dt =$$

$$= \lim_{\lambda\to 0+} \frac{1}{\lambda} \int\limits_{\tau-\lambda}^{\tau} (f(t,\, \hat{x}(t),\, v) - f(t,\, \hat{x}(t),\, \hat{u}(t)))\, dt +$$

$$+ \lim_{\lambda\to 0+} \frac{1}{\lambda} \int\limits_{\tau-\lambda}^{\tau} (f(t,\, x_\lambda(t),\, v) - f(t,\, \hat{x}(t),\, v))\, dt +$$

$$+ \lim_{\lambda\to 0+} \int\limits_{\tau}^{t_1} \frac{1}{\lambda}(f(t,\, x_\lambda(t),\, \hat{u}(t)) - f(t,\, \hat{x}(t),\, \hat{u}(t)))\, dt =: S.$$

Первый предел равен $f(\tau,\, \hat{x}(\tau),\, v) - f(\tau,\, \hat{x}(\tau),\, \hat{u}(\tau))$, второй предел равен 0. Вычислим третий предел. У подынтегрального выражения предел при $\lambda \to 0+$ равен

$$f_x(t,\, \hat{x}(t),\, \hat{u}(t))\, \frac{dx_\lambda(t)}{d\lambda}\bigg|_{\lambda=0} \overset{(62)}{=} f_x(t,\, \hat{x}(t),\, \hat{u}(t)) y(t),$$

где y — решение (63). Так как $x(\cdot,\, \xi)$ липшицево по ξ и существует $C_2 > 0$ такое, что $|\xi_\lambda - \hat{x}(\tau)| \le C_2\lambda$, то существует $A > 0$, не зависящее от t и λ, такое, что

$$\left|\frac{1}{\lambda}(f(t,\, x_\lambda(t),\, \hat{u}(t)) - f(t,\, \hat{x}(t),\, \hat{u}(t)))\right| \le A.$$

Значит, по теореме Лебега о мажорируемой сходимости

$$\lim_{\lambda\to 0+} \int\limits_{\tau}^{t_1} \frac{1}{\lambda}(f(t,\, x_\lambda(t),\, \hat{u}(t)) - f(t,\, \hat{x}(t),\, \hat{u}(t)))\, dt = \int\limits_{\tau}^{t_1} f_x(t,\, \hat{x}(t),\, \hat{u}(t)) y(t)\, dt.$$

Итак,

$$S = f(\tau,\, \hat{x}(\tau),\, v) - f(\tau,\, \hat{x}(\tau),\, \hat{u}(\tau)) + \int\limits_{\tau}^{t_1} f_x(t,\, \hat{x}(t),\, \hat{u}(t)) y(t)\, dt \ge 0. \quad (64)$$

Определим функцию p из уравнения Эйлера (51) и условия трансверсальности (52):

$$\dot p = -p(t)\varphi_x(t,\,\hat x(t),\,\hat u(t)) + f_x(t,\,\hat x(t),\,\hat u(t)),\quad p(t_1)=0. \tag{65}$$

Тогда

$$\int\limits_{\tau}^{t_1} \frac{d(p(t)y(t))}{dt}dt = -p(\tau)y(\tau) \overset{(63)}{=} -p(\tau)(\varphi(\tau,\,\hat x(\tau),\,v) - \varphi(\tau,\,\hat x(\tau),\,\hat u(\tau))). \tag{66}$$

С другой стороны,

$$\int\limits_{\tau}^{t_1} \frac{d(p(t)y(t))}{dt}dt = \int\limits_{\tau}^{t_1} \dot p(t)y(t)\,dt + \int\limits_{\tau}^{t_1} p(t)\dot y(t)\,dt \overset{(63),(65)}{=}$$

$$= \int\limits_{\tau}^{t_1} (-p(t)\varphi_x(t,\,\hat x(t),\,\hat u(t)) + f_x(t,\,\hat x(t),\,\hat u(t)))y(t)\,dt +$$

$$+ \int\limits_{\tau}^{t_1} p(t)\varphi_x(t,\,\hat x(t),\,\hat u(t))y(t)\,dt = \int\limits_{\tau}^{t_1} f_x(t,\,\hat x(t),\,\hat u(t))y(t)\,dt.$$

Из последнего соотношения и (64), (66) получаем, что

$$f(\tau,\,\hat x(\tau),\,v) - f(\tau,\,\hat x(\tau),\,\hat u(\tau)) -$$
$$- p(\tau)(\varphi(\tau,\,\hat x(\tau),\,v) - \varphi(\tau,\,\hat x(\tau),\,\hat u(\tau))) \geq 0,$$

а это и есть неравенство из принципа максимума Понтрягина.

14 Сильный и слабый минимум в простейшей задаче вариационного исчисления. Лемма о скруглении углов.

Дальше будем рассматривать простейшую задачу вариационного исчисления с функциями $x : [t_0,\,t_1] \to \mathbb{R}^n$:

$$\mathcal{L}(x) := \int\limits_{t_0}^{t_1} L(t,\,x,\,\dot x)\,dt \to \inf,\quad x(t_0) = x_{(0)},\ x(t_1) = x_{(1)}. \tag{67}$$

Как и раньше, предполагаем, что L, L_{x_j}, $L_{\dot{x}_j}$ $(1 \leq j \leq n)$ непрерывны.

Задачу (67) можно рассматривать как на пространстве $C^1([t_0, t_1], \mathbb{R}^n)$, так и на пространстве $PC^1([t_0, t_1], \mathbb{R}^n)$.

Определение 4. *Допустимая функция* $\hat{x} \in C^1([t_0, t_1], \mathbb{R}^n)$ *называется точкой слабого локального минимума в задаче* (67), *если существует такое* $\varepsilon > 0$, *что для любой допустимой функции* $x \in C^1([t_0, t_1], \mathbb{R}^n)$ *такой, что* $\|x - \hat{x}\|_{C^1} < \varepsilon$, *выполнено* $\mathcal{L}(x) \geq \mathcal{L}(\hat{x})$.

Определение слабого локального минимума мы уже давали: это локальный минимум относительно C^1-нормы.

Теперь дадим определение точки сильного локального минимума.

Определение 5. *Допустимая функция* $\hat{x} \in PC^1([t_0, t_1], \mathbb{R}^n)$ *называется точкой сильного минимума в задаче* (67), *если существует такое* $\varepsilon > 0$, *что для любых допустимых* $x \in PC^1([t_0, t_1], \mathbb{R}^n)$ *таких, что* $\|x - \hat{x}\|_C < \varepsilon$, *выполнено* $\mathcal{L}(x) \geq \mathcal{L}(\hat{x})$.

То есть сильный минимум — это локальный минимум относительно нормы пространства C. Напомним, что $\|x\|_C \leq \|x\|_{C^1}$. Поэтому, если $\hat{x} \in C^1([t_0, t_1], \mathbb{R}^n)$ — точка сильного минимума, то $\hat{x}$ также будет точкой слабого минимума.

Оказывается, в определении сильного минимума достаточно выполнения неравенства $\mathcal{L}(x) \geq \mathcal{L}(\hat{x})$ для допустимых $x \in C^1([t_0, t_1], \mathbb{R}^n)$, таких что $\|x - \hat{x}\|_C < \varepsilon$. Это вытекает из следующей леммы.

Лемма 6. (о скруглении углов). *Пусть* $x \in PC^1([t_0, t_1], \mathbb{R}^n)$ *— допустимая функция. Тогда для любого* $\varepsilon > 0$ *существует допустимая функция* $x_\varepsilon \in C^1([t_0, t_1], \mathbb{R}^n)$ *такая, что* $\|x_\varepsilon - x\|_C < \varepsilon$ *и* $|\mathcal{L}(x_\varepsilon) - \mathcal{L}(x)| < \varepsilon$.

Доказательство. Пусть $\tau_1 < \dots < \tau_m$ — точки разрыва функции $\dot{x}$,

$$M = \max_{t \in [t_0, t_1]} \max_{1 \leq j \leq n} \max\{|x_j(t)|, |\dot{x}_j(t)|\}.$$

Для малых $\delta > 0$ построим функцию $y_\delta \in C^1([t_0, t_1], \mathbb{R}^n)$ такую, что $y_\delta(t) = x(t)$ при $t \in [t_0, t_1] \setminus \cup_{j=1}^m (\tau_j - \delta, \tau_j + \delta)$, $\|\dot{y}_\delta\|_C \leq M$. Тогда при $t \in (\tau_j - \delta, \tau_j + \delta)$, $1 \leq i \leq n$ выполнено

$$|x_i(t) - y_{\delta,i}(t)| \leq \int\limits_{\tau_j - \delta}^{t} |\dot{x}_i(t) - \dot{y}_{\delta,i}(t)|\, dt \leq 4M\delta.$$

Далее,

$$|\mathcal{L}(y_\delta) - \mathcal{L}(x)| \le \sum_{j=1}^{m} \int\limits_{\tau_j-\delta}^{\tau_j+\delta} |L(t,\, y_\delta(t),\, \dot y_\delta(t)) - L(t,\, x(t),\, \dot x(t))|\, dt \le$$

$$\le 2m\delta \max_{t\in[t_0,\, t_1]} |L(t,\, y_\delta(t),\, \dot y_\delta(t)) - L(t,\, x(t),\, \dot x(t))| \le 2m\delta \cdot C(M,\, L)$$

(здесь $0 < C(M,\, L) < +\infty$; мы воспользовались тем, что L непрерывна и поэтому ограничена на компактах).

Перейдем к построению y_δ. Пусть $z(t) = \dot x(t)$. Нам достаточно для каждого $j = 1,\, \dots,\, m$ построить непрерывную функцию

$$z_{j,\,\delta} = (z_{j,\,\delta,\,1},\, \dots,\, z_{j,\,\delta,\,n}) \in C([\tau_j - \delta,\, \tau_j + \delta],\, \mathbb{R}^n)$$

так, чтобы $\|z_{j,\,\delta}\|_C \le M$, $z_{j,\,\delta}(\tau_j \pm \delta) = z(\tau_j \pm \delta)$, $\int\limits_{\tau_j-\delta}^{\tau_j+\delta} z_{j,\,\delta,\,i}(t)\, dt = \int\limits_{\tau_j-\delta}^{\tau_j+\delta} z_i(t)\, dt$, $1 \le i \le n$.

Пусть для определенности $z_i(\tau_j - 0) < z_i(\tau_j + 0)$.

Для $\tau_j - \delta \le \tau \le \tau_j \le \sigma \le \tau_j + \delta$, $\tau < \sigma$, положим

$$\xi_{\tau,\sigma,i}(t) = \begin{cases} z_i(t),\ t \in [\tau_j - \delta,\, \tau_j + \delta] \backslash (\tau,\, \sigma), \\ z_i(\tau - 0)\frac{\sigma - t}{\sigma - \tau} + z_i(\sigma + 0)\frac{t - \tau}{\sigma - \tau}, \quad t \in [\tau,\, \sigma]. \end{cases}$$

Тогда $\|\xi_{\tau,\sigma,i}\|_C \le M$, функция $\xi_{\tau,\sigma,i}$ непрерывна.

Если δ достаточно мало, то $\int\limits_{\tau_j-\delta}^{\tau_j+\delta} \xi_{\tau,\,\tau_j,\,i}(t)\, dt > \int\limits_{\tau_j-\delta}^{\tau_j+\delta} z_i(t)\, dt$, $\int\limits_{\tau_j-\delta}^{\tau_j+\delta} \xi_{\tau_j,\sigma,i}(t)\, dt < \int\limits_{\tau_j-\delta}^{\tau_j+\delta} z_i(t)\, dt$. Отсюда и из теоремы о промежуточном значении следует, что найдутся такие $\tau < \tau_j < \sigma$, что $\int\limits_{\tau_j-\delta}^{\tau_j+\delta} \xi_{\tau,\sigma,i}(t)\, dt = \int\limits_{\tau_j-\delta}^{\tau_j+\delta} z_i(t)\, dt$. В качестве $z_{j,\delta,i}$ берется $\xi_{\tau,\sigma,i}$. $\square$

Замечание. Для построенной функции y_δ, помимо утверждения леммы, выполнено $\|\dot y_\delta\|_C \le \|\dot x\|_C$.

15 Необходимые условия сильного минимума

Пусть $L : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, L, L_{x_j}, $L_{\dot{x}_j}$ $(1 \le j \le n)$ непрерывны.

Определение 6. *Функция Вейерштрасса $\mathcal{E} : [t_0,\, t_1] \times \mathbb{R}^{3n} \to \mathbb{R}$ определяется по формуле*

$$\mathcal{E}(t,\, x,\, u,\, v) = L(t,\, x,\, v) - L(t,\, x,\, u) - L_{\dot{x}}(t,\, x,\, u)(v - u).$$

Теорема 18. *Пусть $\hat{x} \in PC^1([t_0,\, t_1],\, \mathbb{R}^n)$ — точка сильного минимума в задаче* (67). *Тогда функция $t \mapsto L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ непрерывна на отрезке $[t_0,\, t_1]$; если $\dot{\hat{x}}$ непрерывна в точке t, то $\mathcal{E}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t),\, v) \ge 0$ для любого $v \in \mathbb{R}^n$.*

Доказательство. Сначала заметим, что на интервалах непрерывности функции $\dot{\hat{x}}$ 1) функция $t \mapsto L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ абсолютно непрерывна, 2) выполнено уравнение Эйлера–Лагранжа.

Также отметим, что неравенство $\mathcal{E}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t),\, v) \ge 0$ достаточно проверить для $t \in (t_0,\, t_1)$.

Итак, пусть $\tau \in (t_0,\, t_1)$. Выберем $\sigma \in (\tau,\, t_1]$ так, чтобы функция $\dot{\hat{x}}$ была непрерывна на $(\tau,\, \sigma]$. Также зафиксируем $w \in \mathbb{R}^n$.

Для малых $\lambda > 0$ положим

$$\xi_\lambda(t) = \begin{cases} w, \quad t \in (\tau - \lambda,\, \tau), \\ -\frac{\lambda w}{\sigma - \tau}, \quad t \in (\tau,\, \sigma), \\ 0, \quad \text{иначе}, \end{cases}$$

$h_\lambda(t) = \int\limits_{t_0}^{t} \xi_\lambda(s)\, ds$. Тогда $h_\lambda|_{[t_0,\, \tau-\lambda]\cup[\sigma,\, t_1]} = 0$, $\|h_\lambda\|_C \underset{\lambda \to +0}{\to} 0$. Значит, при малых $\lambda > 0$ выполнено $\mathcal{L}(\hat{x} + h_\lambda) - \mathcal{L}(\hat{x}) \ge 0$. Поэтому

$$\lim_{\lambda \to +0} \frac{\mathcal{L}(\hat{x} + h_\lambda) - \mathcal{L}(\hat{x})}{\lambda} \ge 0. \tag{68}$$

Вычислим этот предел:

$$\lim_{\lambda \to +0} \frac{\mathcal{L}(\hat{x} + h_\lambda) - \mathcal{L}(\hat{x})}{\lambda} =$$

$$= \lim_{\lambda\to+0} \frac{1}{\lambda} \int\limits_{\tau-\lambda}^{\tau} [L(t,\, \hat{x}(t) + h_\lambda(t),\, \dot{\hat{x}}(t) + w) - L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))]\, dt +$$

$$+ \lim_{\lambda\to+0} \frac{1}{\lambda} \int\limits_{\tau}^{\sigma} [L(t,\, \hat{x}(t) + h_\lambda(t),\, \dot{\hat{x}}(t) + \dot{h}_\lambda(t)) - L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))]\, dt.$$

Первый предел равен $L(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau - 0) + w) - L(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau - 0))$ (доказывается так же, как в задаче оптимального управления).

Вычислим второй предел. Сначала заметим, что $\frac{d}{d\lambda} h_\lambda(t)\big|_{\lambda=0} = \frac{w(\sigma-t)}{\sigma-\tau} =: z(t)$, $\frac{d}{d\lambda}\dot{h}_\lambda(t)\Big|_{\lambda=0} = -\frac{w}{\sigma-\tau} = \dot{z}(t)$. Применяя теорему Лагранжа и теорему Лебега о мажорируемой сходимости, получаем, что второй предел равен

$$\int\limits_{\tau}^{\sigma} (L_x(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) z(t) + L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) \dot{z}(t))\, dt.$$

Теперь применим уравнение Эйлера–Лагранжа (на $(\tau,\, \sigma)$ оно выполнено, так как $\dot{\hat{x}}$ непрерывна на этом интервале), затем формулу Ньютона–Лейбница, учтём, что $z(\sigma) = 0$, $z(\tau) = w$, и получим

$$\int\limits_{\tau}^{\sigma} \left(\frac{d}{dt} L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) z(t) + L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) \dot{z}(t) \right) dt =$$

$$= (L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) z(t))\big|_{\tau+0}^{\sigma} = -L_{\dot{x}}(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau + 0)) w.$$

Таким образом, из (68) следует неравенство

$$\begin{gathered} f(w) := L(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau - 0) + w) - L(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau - 0)) - \\ - L_{\dot{x}}(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau + 0)) w \geq 0, \quad w \in \mathbb{R}. \end{gathered} \tag{69}$$

Тогда $f'(0) = 0$, то есть

$$L_{\dot{x}}(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau - 0)) - L_{\dot{x}}(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau + 0)) = 0;$$

это и означает непрерывность функции $t \mapsto L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ в точке τ. Наконец, из (69) следует неравенство $\mathcal{E}(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau),\, v) \geq 0$, $v \in \mathbb{R}^n$, если $\dot{\hat{x}}$ непрерывна в точке τ. $\square$

Задача 21. Сделав замену $\dot{x} = u$, вывести эту теорему из принципа максимума Понтрягина.

Замечание. Для непрерывности функции $t \mapsto L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ достаточно слабого минимума, т.е. локального минимума относительно нормы $\max\{\|x\|_C,\, \|\dot{x}\|_{L_\infty}\}$; см. замечание в §1 про необходимое условие локального минимума на классе липшицевых функций. Это выводится и из (69), так как достаточно рассмотреть только векторы w малой длины и снова написать равенство $f'(0) = 0$.

Определение 7. *Пусть $\hat{x} \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$ — допустимая экстремаль (то есть решение уравнения Эйлера–Лагранжа). Скажем, что выполнено условие Вейерштрасса для $\hat{x}$, если $\mathcal{E}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t),\, v) \geq 0$ для любого $t \in [t_0,\, t_1]$, $v \in \mathbb{R}^n$.*

Таким образом, условие Вейерштрасса является необходимым условием сильного минимума.

Геометрический смысл условия Вейерштрасса следующий. Для каждого $t \in [t_0,\, t_1]$ положим

$$f_t(v) = L(t,\, \hat{x}(t),\, v).$$

Неравенство $\mathcal{E}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t),\, v) \geq 0$ эквивалентно соотношению

$$f_t(v) \geq f_t(\dot{\hat{x}}(t)) + (f_t)'(\dot{\hat{x}}(t))(v - \dot{\hat{x}}(t)).$$

То есть условие Вейерштрасса означает, что для любого $t \in [t_0,\, t_1]$ график функции f_t лежит не ниже касательной гиперплоскости, проведенной в точке $\dot{\hat{x}}(t)$.

Задача 22. Показать, что если L явно не зависит от x (т.е. $L = L(t,\, \dot{x}(t))$), то условие Вейерштрасса будет достаточным условием глобального минимума.

16 Поле экстремалей и функция действия

Пусть $L : [t_0 - \delta,\, t_1 + \delta] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ — непрерывная функция, L_{x_j} и $L_{\dot{x}_j}$ $(1 \leq j \leq n)$ также непрерывны. Как и раньше, рассматриваем задачу

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt \to \inf, \quad x(t_0) = x_{(0)},\ x(t_1) = x_{(1)},\ x \in C^1([t_0,\, t_1],\, \mathbb{R}^n). \tag{70}$$

Определение 8. *Пусть $\hat{x}$ — допустимая экстремаль для задачи (70), $k \in \mathbb{N}$. Скажем, что $\hat{x}$ погружена в C^k-гладкое центральное поле экстремалей, если существуют окрестность V графика $\hat{x}$ и $\alpha_0 \in \mathbb{R}^n$, $\varepsilon > 0$ со следующими свойствами:*

1. *Для каждого $\alpha \in O_\varepsilon(\alpha_0)$ заданы интервал $I_\alpha \subset (t_0 - \delta,\, t_1 + \delta)$ и экстремаль $x(\cdot,\, \alpha) : I_\alpha \to \mathbb{R}^n$, т.е. решение уравнения Эйлера–Лагранжа*
$$-\frac{d}{dt}L_{\dot{x}} + L_x = 0; \tag{71}$$
2. *$x(\cdot,\, \alpha_0) = \hat{x}(\cdot)$;*
3. *отображение $(t,\, \alpha) \mapsto x(t,\, \alpha)$ гладкое порядка C^k;*
4. *график $x(\cdot,\, \alpha)$ лежит в V;*
5. *для любого $(\tau,\, \xi) \in V$ существует единственное $\alpha(\tau,\, \xi) \in O_\varepsilon(\alpha_0)$ такое, что $x(\tau,\, \alpha(\tau,\, \xi)) = \xi$;*
6. *отображение $(\tau,\, \xi) \mapsto \alpha(\tau,\, \xi)$ гладкое порядка C^k.*
7. *экстремали $x(\cdot,\, \alpha)$ продолжаются до некоторой точки $t_* < t_0$ с сохранением (71) и свойств 2, 3, при этом $x(t_*,\, \alpha) = x_*$ для любого α.*

Определим функцию наклона поля $u : V \to \mathbb{R}^n$. Пусть $(\tau,\, \xi) \in V$. Положим $u(\tau,\, \xi) = \frac{d}{dt}\big|_{t=\tau} x(t,\, \alpha(\tau,\, \xi))$.

Пример. Пусть $0 < \delta < \pi/2$, $t_0 = \delta$, $t_1 = \pi - \delta$. Рассмотрим задачу
$$\int\limits_{\delta}^{\pi-\delta} (\dot{x}^2 - x^2)\, dt \to \inf, \quad x(\delta) = x(\pi - \delta) = 0.$$

Тогда семейство экстремалей $x(t,\, \alpha) = \alpha \sin t$, $0 < t < \pi$, $\alpha \in \mathbb{R}$, образуют центральное поле экстремалей, содержащее $\hat{x} = 0$. Функция наклона поля определяется следующим образом. Пусть $\alpha(\tau,\, \xi) \sin\tau = \xi$. Тогда $u(\tau,\, \xi) = \alpha(\tau,\, \xi) \cos t|_{t=\tau} = \xi \operatorname{ctg}\tau$.

Пусть $\{x(\cdot,\,\alpha)\}_{\alpha\in O_\varepsilon(\alpha_0)}$ — центральное поле экстремалей, $x(t_*,\,\alpha)=x_*$. Определим функцию действия равенством

$$S(\tau,\,\xi)=\int\limits_{t_*}^{\tau} L(t,\,x(t,\,\alpha(\tau,\,\xi)),\,\dot x(t,\,\alpha(\tau,\,\xi)))\,dt,\quad (\tau,\,\xi)\in V.$$

Вычислим S'_τ и S'_ξ. Для этого сначала определим функции $p(\tau,\,\xi)$ и $H(\tau,\,\xi)$ равенствами

$$\begin{gathered} p(\tau,\,\xi)=L_{\dot x}(\tau,\,\xi,\,u(\tau,\,\xi)),\\ H(\tau,\,\xi)=u(\tau,\,\xi)L_{\dot x}(\tau,\,\xi,\,u(\tau,\,\xi))-L(\tau,\,\xi,\,u(\tau,\,\xi)) \end{gathered} \tag{72}$$

(для запоминания формул: пишем $p=L_{\dot x}$, $H=\dot xL_{\dot x}-L$, затем подставляем $t=\tau$, $x=\xi$, $\dot x=u(\tau,\,\xi)$). Выражения для p и H у нас были, когда мы изучали законы сохранения.

Предложение 25. *Пусть поле экстремалей имеет гладкость C^2. Тогда выполнены равенства $S'_\tau(\tau,\,\xi)=-H(\tau,\,\xi)$, $S'_\xi(\tau,\,\xi)=p(\tau,\,\xi)$.*

Доказательство. Продифференцировав равенства $x(t_*,\,\alpha(\tau,\,\xi))=x_*$, $x(\tau,\,\alpha(\tau,\,\xi))=\xi$ по τ и по ξ_j ($1\le j\le n$), получим

$$\begin{gathered} \tfrac{\partial}{\partial\tau}\Big|_{t=t_*}x(t,\,\alpha(\tau,\,\xi))=0,\quad \tfrac{\partial}{\partial\tau}\Big|_{t=\tau}x(t,\,\alpha(\tau,\,\xi))=-u(\tau,\,\xi),\\ \tfrac{\partial}{\partial\xi_j}\Big|_{t=t_*}x(t,\,\alpha(\tau,\,\xi))=0,\quad \tfrac{\partial}{\partial\xi_j}\Big|_{t=\tau}x(t,\,\alpha(\tau,\,\xi))=e_j, \end{gathered} \tag{73}$$

где $e_j=(0,\,\dots,\,0,\,1,\,0,\,\dots,\,0)\in\mathbb{R}^n$ (1 стоит на j-м месте). Далее,

$$S'_\tau(\tau,\,\xi)=L(\tau,\,x(\tau,\,\alpha(\tau,\,\xi)),\,\dot x(\tau,\,\alpha(\tau,\,\xi)))+$$

$$+\int\limits_{t_*}^{\tau}\left(L_{\dot x}(t,\,x(t,\,\alpha(\tau,\,\xi)),\,\dot x(t,\,\alpha(\tau,\,\xi)))\frac{\partial}{\partial\tau}\dot x(t,\,\alpha(\tau,\,\xi))+\right.$$

$$\left.+L_x(t,\,x(t,\,\alpha(\tau,\,\xi)),\,\dot x(t,\,\alpha(\tau,\,\xi)))\frac{\partial}{\partial\tau}x(t,\,\alpha(\tau,\,\xi))\right)\,dt\overset{(71)}{=}$$

$$=L(\tau,\,\xi,\,u(\tau,\,\xi))+\int\limits_{t_*}^{\tau}\left(L_{\dot x}(t,\,x(t,\,\alpha(\tau,\,\xi)),\,\dot x(t,\,\alpha(\tau,\,\xi)))\frac{\partial}{\partial t}\frac{\partial}{\partial\tau}x(t,\,\alpha(\tau,\,\xi))+\right.$$

$$+\frac{d}{dt}L_{\dot{x}}(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi)))\frac{\partial}{\partial\tau}x(t,\, \alpha(\tau,\, \xi))\Big)\, dt =$$

$$= L(\tau,\, \xi,\, u(\tau,\, \xi)) + L_{\dot{x}}(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi)))\frac{\partial}{\partial\tau}x(t,\, \alpha(\tau,\, \xi))\Big|_{t_*}^{\tau} \overset{(73)}{=}$$

$$= L(\tau,\, \xi,\, u(\tau,\, \xi)) - L_{\dot{x}}(\tau,\, \xi,\, u(\tau,\, \xi))u(\tau,\, \xi) = -H(\tau,\, \xi),$$

$$S'_{\xi_j}(\tau,\, \xi) = \int\limits_{t_*}^{\tau}\Big(L_{\dot{x}}(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi)))\frac{\partial}{\partial\xi_j}\dot{x}(t,\, \alpha(\tau,\, \xi)) +$$

$$+ L_x(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi)))\frac{\partial}{\partial\xi_j}x(t,\, \alpha(\tau,\, \xi))\Big)\, dt \overset{(71)}{=}$$

$$= \int\limits_{t_*}^{\tau}\Big(L_{\dot{x}}(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi)))\frac{\partial}{\partial t}\frac{\partial}{\partial\xi_j}x(t,\, \alpha(\tau,\, \xi)) +$$

$$+\frac{d}{dt}L_{\dot{x}}(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi)))\frac{\partial}{\partial\xi_j}x(t,\, \alpha(\tau,\, \xi))\Big)\, dt =$$

$$= L_{\dot{x}}(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi)))\frac{\partial}{\partial\xi_j}x(t,\, \alpha(\tau,\, \xi))\Big|_{t_*}^{\tau} \overset{(73)}{=}$$

$$= L_{\dot{x}_j}(\tau,\, \xi,\, u(\tau,\, \xi)) = p_j(\tau,\, \xi).$$

Равенство доказано. □

Таким образом,

$$dS(\tau,\, \xi) = p(\tau,\, \xi)\, d\xi - H(\tau,\, \xi)\, d\tau. \tag{74}$$

17 Формула Вейерштрасса и достаточное условие сильного экстремума

Напомним, что функция Вейерштрасса задается формулой

$$\mathcal{E}(t,\, x,\, u,\, v) = L(t,\, x,\, v) - L(t,\, x,\, u) - L_u(t,\, x,\, u)(v-u).$$

Теорема 19. *Пусть $\hat{x} \in C^2([t_0, t_1], \mathbb{R}^n)$ — допустимая экстремаль для задачи* (70), *окруженная C^2-гладким центральным полем экстремалей, $x \in C^1([t_0, t_1], \mathbb{R}^n)$ — допустимая функция, и пусть график x лежит во множестве V из определения* 8. *Тогда*

$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt - \int_{t_0}^{t_1} L(t, \hat{x}(t), \dot{\hat{x}}(t))\,dt = \int_{t_0}^{t_1} \mathcal{E}(t, x(t), u(t, x(t)), \dot{x}(t))\,dt. \tag{75}$$

Доказательство. Пусть $\gamma(t) = (t, x(t))$, $\hat{\gamma}(t) = (t, \hat{x}(t))$, $t \in [t_0, t_1]$. Тогда $\int\limits_\gamma dS = \int\limits_{\hat{\gamma}} dS$; в силу (74), $\int\limits_\gamma (p\,dx - H\,dt) = \int\limits_{\hat{\gamma}} (p\,dx - H\,dt)$. Значит,

$$\int_{t_0}^{t_1} (L_{\dot{x}}(t, x(t), u(t, x(t)))\dot{x}(t) - u(t, x(t))L_{\dot{x}}(t, x(t), u(t, x(t)))+$$

$$+L(t, x(t), u(t, x(t))))\,dt =$$

$$= \int_{t_0}^{t_1} (L_{\dot{x}}(t, \hat{x}(t), u(t, \hat{x}(t)))\dot{\hat{x}}(t) - u(t, \hat{x}(t))L_{\dot{x}}(t, \hat{x}(t), u(t, \hat{x}(t)))+$$

$$+L(t, \hat{x}(t), u(t, \hat{x}(t))))\,dt,$$

т.е.

$$\int_{t_0}^{t_1} (L_{\dot{x}}(t, x(t), u(t, x(t)))(\dot{x}(t) - u(t, x(t))) + L(t, x(t), u(t, x(t))))\,dt =$$

$$= \int_{t_0}^{t_1} L(t, \hat{x}(t), \dot{\hat{x}}(t)))\,dt.$$

Вычтем обе части равенства из $\int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t)))\,dt$ и получим (75). $\square$

Определение 9. *Пусть $\hat{x}$ — допустимая экстремаль. Скажем, что выполнено усиленное условие Вейерштрасса, если существует $\varepsilon > 0$ такое, что для любого $t \in [t_0, t_1]$, для любых $x \in \mathbb{R}^n$, $u \in \mathbb{R}^n$ таких, что $|x - \hat{x}(t)| < \varepsilon$, $|u - \dot{\hat{x}}(t)| < \varepsilon$, и для любого $v \in \mathbb{R}^n$ выполнено $\mathcal{E}(t, x, u, v) \geq 0$.*

Теорема 20. *Пусть $\hat{x}$ — допустимая экстремаль, погруженная в C^2-гладкое центральное поле экстремалей. Предположим, что выполнено усиленное условие Вейерштрасса. Тогда $\hat{x}$ — точка сильного минимума.*

Доказательство. Если $\|x-\hat{x}\|_C$ достаточно мало, то график x лежит во множестве V из определения 8. Значит, выполнена формула Вейерштрасса.

Кроме того, если $\|x-\hat{x}\|_C$ мало, то $|x(t)-\hat{x}(t)| < \varepsilon$, $|u(t,\, x(t))-\dot{\hat{x}}(t)| < \varepsilon$, где $\varepsilon > 0$ — из определения усиленного условия Вейерштрасса. Значит, для любого $t \in [t_0,\, t_1]$ выполнено $\mathcal{E}(t,\, x(t),\, u(t,\, x(t)),\, \dot{x}(t)) \geq 0$, так что

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt - \int\limits_{t_0}^{t_1} L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))\, dt =$$

$$= \int\limits_{t_0}^{t_1} \mathcal{E}(t,\, x(t),\, u(t,\, x(t)),\, \dot{x}(t))\, dt \geq 0.$$

Теорема доказана. □

Замечание. Геометрический смысл усиленного условия Вейерштрасса следующий. Рассмотрим для фиксированных x, t функцию $f_{t,x}(v) = L(t,\, x,\, v)$. Усиленное условие Вейерштрасса выполнено, если существует такое $\varepsilon > 0$, что для любого $t \in [t_0,\, t_1]$, для любого $x \in O_\varepsilon(\hat{x}(t))$, для любого $u \in O_\varepsilon(\dot{\hat{x}}(t))$ график функции $f_{t,x}$ лежит не ниже касательной гиперплоскости, проведенной в точке u.

Замечание. Если L выпукла по $\dot{x}$, то $\mathcal{E}(t,\, x,\, u,\, v) \geq 0$ для любых x, u, $v \in \mathbb{R}^n$, $t \in [t_0,\, t_1]$. Пусть также в качестве V из определения 8 можно взять $(t_0-\delta,\, t_1+\delta)\times\mathbb{R}^n$. Тогда в рассуждениях выше допустимая функция x может быть произвольной (а не из окрестности $\hat{x}$ относительно нормы C). В этом случае $\hat{x}$ — точка глобального минимума.

Задача 23. Доказать, используя поле экстремалей, что допустимая экстремаль в задаче

$$\int\limits_{t_0}^{t_1} \frac{\sqrt{1+\dot{x}^2}}{x}\, dt \to \mathrm{extr}, \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad x > 0$$

является точкой глобального минимума (здесь $x_0 > 0$, $x_1 > 0$).

Задача 24. Доказать, используя поле экстремалей, что допустимая экстремаль в задаче

$$\int\limits_{t_0}^{t_1} \frac{\sqrt{1+\dot{x}^2}}{\sqrt{x}}\, dt \to \mathrm{extr}, \quad x(t_0) = x_0, \quad x(t_1) = x_1, \quad x > 0$$

является точкой глобального минимума (здесь $x_0 > 0$, $x_1 > 0$).

Указание. Нужно сначала проверить, что для любых t_*, $\tau \in \mathbb{R}$, $x_* > 0$, $\xi > 0$ существует экстремаль x с граничными условиями $x(t_*) = x_*$, $x(\tau) = \xi$, причем она гладко зависит от τ и ξ. В задаче о геодезических на плоскости Лобачевского это легко сделать из геометрических соображений. В задаче о брахистохроне нужно воспользоваться тем, что решения уравнения Эйлера получаются друг из друга с помощью гомотетии и сдвига, что их производные строго убывают и принимают все вещественные значения.

18 Необходимые условия слабого минимума

Далее будем дополнительно предполагать, что существуют вторые частные производные $L_{\dot{x}_i\dot{x}_j}$, $L_{\dot{x}_i x_j}$, $L_{x_i\dot{x}_j}$, $L_{x_i x_j}$ $(1 \le i,\, j \le n)$, при этом отображения $(t,\, \xi,\, \eta) \mapsto L_{\dot{x}_i\dot{x}_j}(t,\, \xi,\, \eta)$, $(t,\, \xi,\, \eta) \mapsto L_{\dot{x}_i x_j}(t,\, \xi,\, \eta)$, $(t,\, \xi,\, \eta) \mapsto L_{x_i\dot{x}_j}(t,\, \xi,\, \eta)$, $(t,\, \xi,\, \eta) \mapsto L_{x_i x_j}(t,\, \xi,\, \eta)$ непрерывны.

Замечание. В силу непрерывности вторых частных производных, $L_{\dot{x}_i\dot{x}_j} = L_{\dot{x}_j\dot{x}_i}$, $L_{\dot{x}_i x_j} = L_{x_j\dot{x}_i}$, $L_{x_i x_j} = L_{x_j x_i}$.

Пусть $\hat{x}$ — допустимая экстремаль. Будем обозначать

$$\hat{L}_{x_j}(t) = L_{x_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)),\ \hat{L}_{\dot{x}_j}(t) = L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)),$$

$$\hat{L}_{x_i x_j}(t) = L_{x_i x_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)),\ \hat{L}_{\dot{x}_i x_j}(t) = L_{\dot{x}_i x_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)),$$

$$\hat{L}_{\dot{x}_i\dot{x}_j}(t) = L_{\dot{x}_i\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)).$$

Пусть $h \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$, $h(t_0) = h(t_1) = 0$. Применяя формулу Тейлора, получаем, что для каждого $\lambda \in \mathbb{R}$, $t \in [t_0,\, t_1]$ существует $\theta_{\lambda,t} \in [0,\, 1]$

такое, что

$$\mathcal{L}(\hat{x}+\lambda h)-\mathcal{L}(\hat{x})=\lambda\int\limits_{t_0}^{t_1}\left(\sum_{j=1}^{n}(\hat{L}_{\dot{x}_j}(t)\dot{h}_j(t)+\hat{L}_{x_j}(t)h_j(t))\right)dt+$$

$$+\frac{\lambda^2}{2}\int\limits_{t_0}^{t_1}\sum_{i,j=1}^{n}L_{\dot{x}_i\dot{x}_j}(t,\,\hat{x}(t)+\lambda\theta_{\lambda,t}h(t),\,\dot{\hat{x}}(t)+\lambda\theta_{\lambda,t}\dot{h}(t))\dot{h}_i(t)\dot{h}_j(t)\,dt+$$

$$+\frac{\lambda^2}{2}\int\limits_{t_0}^{t_1}2\sum_{i,j=1}^{n}L_{\dot{x}_ix_j}(t,\,\hat{x}(t)+\lambda\theta_{\lambda,t}h(t),\,\dot{\hat{x}}(t)+\lambda\theta_{\lambda,t}\dot{h}(t))\dot{h}_i(t)h_j(t)\,dt+$$

$$+\frac{\lambda^2}{2}\int\limits_{t_0}^{t_1}\sum_{i,j=1}^{n}L_{x_ix_j}(t,\,\hat{x}(t)+\lambda\theta_{\lambda,t}h(t),\,\dot{\hat{x}}(t)+\lambda\theta_{\lambda,t}\dot{h}(t))h_i(t)h_j(t)\,dt.$$

Так как $\hat{x}$ — допустимая экстремаль, то первое слагаемое равно 0. Далее, отображения $(t,\,\xi,\,\eta)\mapsto L_{\dot{x}_i\dot{x}_j}(t,\,\xi,\,\eta)$, $(t,\,\xi,\,\eta)\mapsto L_{\dot{x}_ix_j}(t,\,\xi,\,\eta)$, $(t,\,\xi,\,\eta)\mapsto L_{x_ix_j}(t,\,\xi,\,\eta)$ равномерно непрерывны на компактах, поэтому для любого $h\in C^1([t_0,\,t_1],\,\mathbb{R}^n)$

$$\begin{aligned}
&L_{\dot{x}_i\dot{x}_j}(t,\,\hat{x}(t)+\lambda\theta_{\lambda,t}h(t),\,\dot{\hat{x}}(t)+\lambda\theta_{\lambda,t}\dot{h}(t))\underset{[t_0,\,t_1]}{\rightrightarrows}\hat{L}_{\dot{x}_i\dot{x}_j}(t),\quad\lambda\to0,\\
&L_{\dot{x}_ix_j}(t,\,\hat{x}(t)+\lambda\theta_{\lambda,t}h(t),\,\dot{\hat{x}}(t)+\lambda\theta_{\lambda,t}\dot{h}(t))\underset{[t_0,\,t_1]}{\rightrightarrows}\hat{L}_{\dot{x}_ix_j}(t),\quad\lambda\to0,\\
&L_{x_ix_j}(t,\,\hat{x}(t)+\lambda\theta_{\lambda,t}h(t),\,\dot{\hat{x}}(t)+\lambda\theta_{\lambda,t}\dot{h}(t))\underset{[t_0,\,t_1]}{\rightrightarrows}\hat{L}_{x_ix_j}(t),\quad\lambda\to0.
\end{aligned}\tag{76}$$

Предложение 26. *Пусть $\hat{x}$ — точка слабого минимума. Тогда для любого $h\in C^1([t_0,\,t_1],\,\mathbb{R}^n)$ такого, что $h(t_0)=h(t_1)=0$, выполнено*

$$\begin{aligned}
\int\limits_{t_0}^{t_1}\Bigg(\sum_{i,j=1}^{n}\hat{L}_{\dot{x}_i\dot{x}_j}(t)\dot{h}_i(t)\dot{h}_j(t)+2\sum_{i,j=1}^{n}\hat{L}_{\dot{x}_ix_j}(t)\dot{h}_i(t)h_j(t)+\\
+\sum_{i,j=1}^{n}\hat{L}_{x_ix_j}(t)h_i(t)h_j(t)\Bigg)\,dt\geq0.
\end{aligned}\tag{77}$$

Доказательство. Так как $\hat{x}$ — точка слабого минимума, она удовлетворяет уравнению Эйлера–Лагранжа. Значит, в силу (76),

$$\mathcal{L}(\hat{x}+\lambda h)-\mathcal{L}(\hat{x})=\frac{\lambda^2}{2}\int\limits_{t_0}^{t_1}\left(\sum_{i,j=1}^{n}\hat{L}_{\dot{x}_i\dot{x}_j}(t)\dot{h}_i(t)\dot{h}_j(t)+\right.$$

$$+2\sum_{i,j=1}^{n}\hat{L}_{\dot{x}_i x_j}(t)\dot{h}_i(t)h_j(t)+\sum_{i,j=1}^{n}\hat{L}_{x_i x_j}(t)h_i(t)h_j(t)\Bigg)\,dt+\underset{\lambda\to 0}{o(\lambda^2)}.$$

При малых λ левая часть неотрицательна; отсюда следует (77). $\square$

Через $\hat{L}_{\dot{x}\dot{x}}(t)$, $\hat{L}_{\dot{x}x}(t)$, $\hat{L}_{xx}(t)$ будем обозначать матрицу, у которой на пересечении i-й строки и j-го столбца стоит соответственно $\hat{L}_{\dot{x}_i\dot{x}_j}(t)$, $\hat{L}_{\dot{x}_i x_j}(t)$, $\hat{L}_{x_i x_j}(t)$. Матрицы $\hat{L}_{\dot{x}\dot{x}}(t)$ и $\hat{L}_{xx}(t)$ симметричные.

Далее для квадратной $n\times n$ матрицы A и n-мерных векторов v, w будем через $w\cdot Av$ обозначать скалярное произведение векторов w и Av (A умножается на вектор-столбец v).

Таким образом, левая часть (77) записывается как

$$\int_{t_0}^{t_1}(\dot{h}\cdot\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}+2\dot{h}\cdot\hat{L}_{\dot{x}x}(t)h+h\cdot\hat{L}_{xx}(t)h)\,dt.$$

18.1 Условие Лежандра

Пусть A — симметричная матрица $n\times n$. Будем писать $A\geq 0$ ($A>0$), если A — неотрицательно определенная (соответственно положительно определенная) матрица.

Пусть $\hat{x}$ — допустимая экстремаль. Скажем, что выполнено условие Лежандра, если $\hat{L}_{\dot{x}\dot{x}}(t)\geq 0$ для любого $t\in[t_0,\,t_1]$; выполнено усиленное условие Лежандра, если $\hat{L}_{\dot{x}\dot{x}}(t)>0$ для любого $t\in[t_0,\,t_1]$.

Теорема 21. *Пусть $\hat{x}$ — точка слабого минимума. Тогда выполнено условие Лежандра.*

Доказательство. Ранее было показано, что для любой функции $h\in C^1([t_0,\,t_1],\,\mathbb{R}^n)$, такой, что $h(t_0)=h(t_1)=0$, выполнено (77).

Пусть условие Лежандра не выполнено. Тогда существует точка $\tau\in[t_0,\,t_1]$ и вектор $w\in\mathbb{R}^n$ такие, что $w\cdot\hat{L}_{\dot{x}\dot{x}}(\tau)w<0$ и $\|w\|=1$ (евклидова длина). Без ограничения общности можно считать, что $\tau\in(t_0,\,t_1)$. Тогда существуют $a>0$, $M>0$ такие, что в некоторой окрестности точки τ выполнено $w\cdot\hat{L}_{\dot{x}\dot{x}}(t)w<-a$, $\|\hat{L}_{\dot{x}x}(t)\|\leq M$, $\|\hat{L}_{xx}(t)\|\leq M$ (здесь $\|\cdot\|$ — норма оператора, действующего на $\mathbb{R}^n$). Для малых $\delta>0$ положим

$$h_\delta(t)=\begin{cases}(\delta-|t-\tau|)w, & t\in[\tau-\delta,\,\tau+\delta],\\ 0, & \text{иначе.}\end{cases}$$

Тогда

$$\int\limits_{t_0}^{t_1} (\dot h_\delta(t) \cdot \hat L_{\dot x\dot x}(t)\dot h_\delta(t) + 2\dot h_\delta(t) \cdot \hat L_{\dot x x}(t)h_\delta(t) + h_\delta(t) \cdot \hat L_{xx}(t)h_\delta(t))\,dt =$$

$$= \int\limits_{\tau-\delta}^{\tau+\delta} (\dot h_\delta(t) \cdot \hat L_{\dot x\dot x}(t)\dot h_\delta(t) + 2\dot h_\delta(t)\hat L_{\dot x x}(t)h_\delta(t) + h_\delta(t) \cdot \hat L_{xx}(t)h_\delta(t))\,dt \le$$

$$\le -\int\limits_{\tau-\delta}^{\tau+\delta} a\,dt + 2M\int\limits_{\tau-\delta}^{\tau+\delta} (\delta - |t-\tau|)\,dt + M\int\limits_{\tau-\delta}^{\tau+\delta} (\delta - |t-\tau|)^2\,dt \le$$

$$\le -2a\delta + 4M\delta^2 + 2M\delta^3 < 0,$$

если $\delta > 0$ достаточно мало.

Построенная функция h_δ кусочно-гладкая. Осталось применить лемму о скруглении углов. □

18.2 Условие Якоби

Обозначим

$$J(h) = \int\limits_{t_0}^{t_1} (\dot h \cdot \hat L_{\dot x\dot x}(t)\dot h + 2\dot h \cdot \hat L_{\dot x x}(t)h + h \cdot \hat L_{xx}(t)h)\,dt.$$

Уравнением Якоби называется уравнение Эйлера–Лагранжа для J. Оно имеет вид:

$$-\frac{d}{dt}(\hat L_{\dot x\dot x}(t)\dot h(t) + \hat L_{\dot x x}(t)h(t)) + (\hat L_{\dot x x}(t))^T \dot h(t) + \hat L_{xx}(t)h(t) = 0, \tag{78}$$

или

$$-\frac{d}{dt}\left(\sum_{j=1}^n \hat L_{\dot x_i\dot x_j}(t)\dot h_j(t) + \sum_{j=1}^n \hat L_{\dot x_i x_j}(t)h_j(t)\right) +$$

$$+\sum_{j=1}^n \hat L_{x_i\dot x_j}(t)\dot h_j(t) + \sum_{j=1}^n \hat L_{x_i x_j}(t)h_j(t) = 0, \quad 1 \le i \le n.$$

Это линейное уравнение. Сделаем замену $w(t) = \hat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t) + \hat{L}_{\dot{x}x}(t)h(t)$. Тогда (78) записывается в виде системы

$$\dot{w} = (\hat{L}_{\dot{x}x}(t))^T\dot{h}(t) + \hat{L}_{xx}(t)h(t), \quad \hat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t) = w - \hat{L}_{\dot{x}x}(t)h(t).$$

Если $\det \hat{L}_{\dot{x}\dot{x}}(t) \neq 0$ для любого $t \in [t_0, t_1]$, то ко второму уравнению можно применить матрицу $(\hat{L}_{\dot{x}\dot{x}}(t))^{-1}$ и получить систему

$$\begin{pmatrix} \dot{h} \\ \dot{w} \end{pmatrix} = C(t) \begin{pmatrix} h \\ w \end{pmatrix},$$

где $C(t)$ — матрица с непрерывными коэффициентами. У этой системы двумерное пространство решений. При этом решение однозначно определяется по значению $h(\tau)$ и $\dot{h}(\tau)$ в некоторой точке $\tau \in [t_0, t_1]$.

Определение 10. *Точка $\tau \in (t_0, t_1]$ называется сопряженной к t_0, если существует ненулевое решение уравнения* (78) *такое, что $h(t_0) = h(\tau) = 0$.*

Определение 11. *Скажем, что выполнено условие Якоби, если на (t_0, t_1) нет сопряженных к t_0 точек; выполнено усиленное условие Якоби, если на $(t_0, t_1]$ нет сопряженных к t_0 точек.*

Теорема 22. *Пусть $\hat{x}$ — точка слабого минимума, при этом выполнено усиленное условие Лежандра. Тогда выполнено условие Якоби.*

Доказательство. Пусть $\tau \in (t_0, t_1)$ — сопряженная точка, $h_0 \neq 0$ — решение уравнения (78), $h_0(t_0) = h_0(\tau) = 0$. Положим $h_*(t) = h_0(t)$ при $0 \le t \le \tau$, $h_*(t) = 0$ при $\tau \le t \le t_1$.

Покажем, что $J(h_*) = 0$. Так как h_0 — решение уравнения (78), то

$$\int_{t_0}^{\tau} h_*(t) \left(-\frac{d}{dt}(\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}_*(t) + \hat{L}_{\dot{x}x}(t)h_*(t)) + (\hat{L}_{\dot{x}x}(t))^T\dot{h}_*(t) + \hat{L}_{xx}(t)h_*(t) \right) dt = 0.$$

Проинтегрировав по частям и учитывая, что $h_*(t_0) = h_*(\tau) = 0$, получаем

$$\int_{t_0}^{\tau} (\dot{h}_* \cdot \hat{L}_{\dot{x}\dot{x}}(t)\dot{h}_* + 2\dot{h}_* \cdot \hat{L}_{\dot{x}x}(t)h_* + h_* \cdot \hat{L}_{xx}(t)h_*)\, dt = 0.$$

Отсюда следует, что $J(h_*) = 0$.

Если $\hat{x}$ — точка слабого минимума, то $J(h) \geq 0$ для любого $h \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$ такого, что $h(t_0) = h(t_1) = 0$. По лемме о скруглении углов, это же неравенство выполнено и для кусочно-гладких h, равных нулю в концах отрезка.

Таким образом, h_* — точка сильного (даже глобального) минимума J. Тогда функция $t \mapsto \hat{L}_{\dot{x}\dot{x}}(t)\dot{h}_*(t) + \hat{L}_{\dot{x}x}(t)h_*(t)$ непрерывна (см. теорему 18). В частности, $\hat{L}_{\dot{x}\dot{x}}(\tau)\dot{h}_*(\tau - 0) = 0$. Так как матрица $\hat{L}_{\dot{x}\dot{x}}(\tau)$ невырожденная, то отсюда $\dot{h}_0(\tau) = 0$. Но тогда по теореме единственности $h_0 = 0$. Противоречие. $\square$

Замечание. Рассмотрим задачу $\int\limits_0^1 \dot{x}^4\, dt \to \inf$, $x(0) = x(1) = 0$. Тогда $\hat{x} = 0$ — точка глобального минимума. При этом $\hat{L}_{\dot{x}\dot{x}} = \hat{L}_{\dot{x}x} = \hat{L}_{xx} = 0$. Значит, выполнено условие Лежандра (неусиленное). При этом уравнение Якоби имеет вид $0 = 0$, т.е. любая функция будет его решением. Значит, условие Якоби не выполнено.

19 Достаточные условия сильного и слабого минимума

Как и раньше, рассматривается задача

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt \to \inf, \quad x(t_0) = x_{(0)},\ x(t_1) = x_{(1)},\ x \in C^1([t_0,\, t_1],\, \mathbb{R}^n). \tag{79}$$

Теорема 23. *Пусть $L \in C^3([t_0 - \delta,\, t_1 + \delta] \times \mathbb{R}^n \times \mathbb{R}^n)$ для некоторого $\delta > 0$. Пусть $\hat{x}$ — допустимая экстремаль в задаче* (79). *Предположим, что выполнены усиленное условие Лежандра и усиленное условие Якоби. Тогда $\hat{x}$ можно погрузить в C^2-гладкое центральное поле экстремалей.*

Эту теорему мы докажем позже, а сейчас получим два следствия из нее.

Теорема 24. *Пусть $L \in C^3([t_0 - \delta,\, t_1 + \delta] \times \mathbb{R}^n \times \mathbb{R}^n)$ для некоторого $\delta > 0$, $\hat{x}$ — допустимая экстремаль в задаче* (79). *Предположим, что выполнены усиленное условие Лежандра и усиленное условие Якоби. Тогда $\hat{x}$ — точка слабого минимума.*

Доказательство. По теореме 23, экстремаль $\hat{x}$ вписывается в C^2-гладкое центральное поле экстремалей.

Пусть x — допустимая функция, $\|x-\hat{x}\|_{C^1} < \varepsilon$. Если $\varepsilon > 0$ достаточно мало, то график функции x лежит во множестве V из определения поля экстремалей. Значит, выполнена формула Вейерштрасса:

$$\mathcal{L}(x) - \mathcal{L}(\hat{x}) = \int\limits_{t_0}^{t_1} \mathcal{E}(t,\, x(t),\, u(t,\, x(t)),\, \dot{x}(t))\, dt.$$

Из формулы Тейлора и определения $\mathcal{E}$ следует, что

$$\mathcal{E}(t,\, x(t),\, u(t,\, x(t)),\, \dot{x}(t)) =$$

$$= \frac{1}{2}(\dot{x}(t) - u(t,\, x(t))) \cdot L_{\dot{x}\dot{x}}(t,\, x(t),\, w(t))(\dot{x}(t) - u(t,\, x(t))),$$

где $w(t) \in [u(t,\, x(t)),\, \dot{x}(t)]$. Если $\varepsilon > 0$ достаточно мало, то $\|x - \hat{x}\|_C$, $\|\dot{x} - \dot{\hat{x}}\|_C$ и $\|u(\cdot,\, x(\cdot)) - \dot{\hat{x}}\|_C$ малы, а значит, коэффициенты матрицы $L_{\dot{x}\dot{x}}(t,\, x(t),\, w(t))$ близки к коэффициентам матрицы $L_{\dot{x}\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$; значит, при малых $\varepsilon > 0$ матрица $L_{\dot{x}\dot{x}}(t,\, x(t),\, w(t))$ неотрицательно определенная (так как выполнено усиленное условие Лежандра) и

$$\mathcal{L}(x){-}\mathcal{L}(\hat{x}) = \int\limits_{t_0}^{t_1} \frac{1}{2}(\dot{x}(t){-}u(t,\, x(t)))\cdot L_{\dot{x}\dot{x}}(t,\, x(t),\, w(t))(\dot{x}(t){-}u(t,\, x(t)))\, dt \geq 0.$$

Теорема доказана. $\square$

Задача 25. Рассмотрим задачу $\int\limits_0^{\pi} (\dot{x}^2 - x^2 - x^4)\, dt \to \inf$, $x(0) = x(\pi) = 0$. Показать, что для $\hat{x} = 0$ выполнено усиленное условие Лежандра, условие Якоби, при этом $\hat{x}$ не является точкой слабого минимума.

Задача 26. Рассмотрим задачу $\int\limits_0^{3/2} (\dot{x}^3 + 2x)\, dt \to \inf$, $x(0) = 0$, $x(3/2) = 1$. Доказать, что выполнено (неусиленное) условие Лежандра, усиленное условие Якоби, но допустимая экстремаль — не точка слабого минимума. (**Указание**. Взять приращения h_ε такие, чтобы $\dot{h}_\varepsilon(t) = -\varepsilon$ при $0 \leq t \leq \delta_\varepsilon$, $\dot{h}_\varepsilon(t) = c_\varepsilon$ при $\delta_\varepsilon \leq t \leq 3/2$, числа c_ε и δ_ε подобрать. Затем воспользоваться леммой о скруглении углов.)

Теорема 25. *Пусть* $L \in C^3([t_0 - \delta,\, t_1 + \delta] \times \mathbb{R}^n \times \mathbb{R}^n)$. *Пусть* $\hat{x}$ *— допустимая экстремаль в задаче* (79). *Предположим, что выполнены усиленное условие Лежандра, усиленное условие Якоби и усиленное условие Вейерштрасса. Тогда* $\hat{x}$ *— точка сильного минимума.*

Доказательство. Утверждение следует из теорем 20 и 23. □

Задача 27. Рассмотрим задачу $\int\limits_0^1 (\dot{x}^2 - x\dot{x}^3)\, dt \to \mathrm{extr}$, $x(0) = x(1) = 0$. Показать, что для экстремали $\hat{x} \equiv 0$ выполнено усиленное условие Лежандра, усиленное условие Якоби, условие Вейерштрасса (не усиленное) и $\hat{x}$ не является точкой сильного минимума.

Задача 28. Рассмотрим задачу

$$\int\limits_{-T_0}^{T_0} x\sqrt{\dot{x}^2 + 1}\, dt \to \min,\ x(T_0) = x(-T_0) = \xi.$$

1) Выписать уравнение Якоби, подобрать одно из его решений, затем найти общее решение. 2) Пусть допустимых экстремалей две. Доказать, что одна из них является точкой сильного минимума, а вторая не является точкой слабого минимума.

Итак, пусть $L \in C^3([t_0 - \delta,\, t_1 + \delta] \times \mathbb{R} \times \mathbb{R})$. Тогда исследование допустимой экстремали проводится следующим образом.

1. Проверяем условие Лежандра.
    - Если оно не выполнено, то нет слабого минимума.
    - Если условие Лежандра выполнено, а усиленное условие Лежандра — нет, то это особый случай и приведенная теория ответа не дает.
    - Если выполнено усиленное условие Лежандра, то переходим к следующему пункту.
2. Проверяем условие Якоби.
    - Если оно не выполнено, то нет слабого минимума.
    - Если условие Якоби выполнено, а усиленное условие Якоби — нет, то это особый случай.

- Если выполнено усиленное условие Якоби, то есть слабый минимум. Переходим к следующему пункту для исследования на сильный минимум.

3. Проверяем условие Вейерштрасса.

   - Если оно не выполнено, то нет сильного минимума.
   - Если условие Вейерштрасса выполнено, а усиленное условие Вейерштрасса — нет, то это особый случай.
   - Если выполнено усиленное условие Вейерштрасса, то есть сильный минимум.

Наша дальнейшая цель — доказать теорему 23.

19.1 Гамильтонова форма уравнений Эйлера

Будем писать $L = L(t,\, x,\, u)$. Положим $\hat{p}(t) = L_u(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$.

По условию теоремы 23, для любого $t \in [t_0,\, t_1]$ матрица $L_{\dot{x}\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ положительно определенная; в частности, у нее ненулевой определитель. По теореме об обратном отображении, существует окрестность точки $(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$, диффеоморфно отображающаяся на окрестность точки $(t,\, \hat{x}(t),\, \hat{p}(t))$ следующим образом:

$$(t,\, x,\, u) \mapsto (t,\, x,\, p) = (t,\, x,\, L_u(t,\, x,\, u)).$$

Обратное отображение имеет вид $(t,\, x,\, p) \mapsto (t,\, x,\, u(t,\, x,\, p))$, где $u(t,\, x,\, p)$ удовлетворяет равенству

$$L_u(t,\, x,\, u(t,\, x,\, p)) = p. \tag{80}$$

Теперь положим

$$H(t,\, x,\, p) = p \cdot u(t,\, x,\, p) - L(t,\, x,\, u(t,\, x,\, p)).$$

Вычислим частные производные функции H по p и по x:

$$H_{x_j}(t,\, x,\, p) = p \cdot u_{x_j}(t,\, x,\, p) - L_{x_j}(t,\, x,\, u(t,\, x,\, p)) -$$

$$-L_u(t,\, x,\, u(t,\, x,\, p))u_{x_j}(t,\, x,\, p) \overset{(80)}{=} -L_{x_j}(t,\, x,\, u(t,\, x,\, p)),$$

$$H_{p_j}(t,\, x,\, p) = u_j(t,\, x,\, p) + p \cdot u_{p_j}(t,\, x,\, p) -$$

$$-L_u(t,\, x,\, u(t,\, x,\, p))u_{p_j}(t,\, x,\, p) \overset{(80)}{=} u_j(t,\, x,\, p).$$

Значит,

$$H_x(t,\, x,\, p) = -L_x(t,\, x,\, u(t,\, x,\, p)), \quad H_p(t,\, x,\, p) = u(t,\, x,\, p). \tag{81}$$

Имеем $\dot{\hat{x}}(t) = u(t,\, \hat{x}(t),\, \hat{p}(t))$; из уравнения Эйлера следует, что $\dot{\hat{p}}(t) = L_x(t,\, x,\, \dot{\hat{x}}(t))$. Отсюда и из (81) получаем

$$\dot{\hat{x}}(t) = H_p(t,\, \hat{x}(t),\, \hat{p}(t)), \quad \dot{\hat{p}}(t) = -H_x(t,\, \hat{x}(t),\, \hat{p}(t)).$$

Функция H определена в окрестности графика $(\hat{x},\, \hat{p})$. Рассмотрим в этой окрестности *систему Гамильтона*:

$$\dot{x} = H_p(t,\, x,\, p), \quad \dot{p} = -H_x(t,\, x,\, p). \tag{82}$$

Пара $(\hat{x},\, \hat{p})$ является одним из ее решений на $[t_0 - \gamma,\, t_1 + \gamma]$, где $\gamma \in (0,\, \delta)$. Пусть $t_* \in [t_0 - \gamma,\, t_0)$, $x_* := \hat{x}(t_*)$, $p_* := \hat{p}(t_*)$. Тогда существует окрестность U точки $(x_*,\, p_*)$ такая, что для любых $(\xi,\, \eta) \in U$ решение (82) с начальными условиями $x(t_*) = \xi$, $p(t_*) = \eta$ продолжается на весь отрезок $[t_*,\, t_1 + \gamma]$. (Одномерный аналог этого утверждения доказывался при исследовании задачи оптимального управления, в многомерном случае доказательство аналогичное.) Кроме того, так как $L \in C^3$, то из формул следует, что $H_x,\, H_p \in C^2$. Отсюда можно вывести, что решение (82) C^2-гладко зависит от $(t,\, \xi,\, \eta)$ (подробное доказательство мы опустим).

Так как мы хотим построить центральное поле экстремалей, то далее будем рассматривать только решения (82), удовлетворяющие условию $x(t_*) = x_*$.

Пусть $x_\eta(\cdot)$, $p_\eta(\cdot)$ — решение (82) с начальными условиями $x(t_*) = x_*$, $p(t_*) = \eta$. Покажем, что $x_\eta(\cdot)$ — решение уравнения Эйлера–Лагранжа. В самом деле, из равенства $H_p(t,\, x,\, p) \overset{(81)}{=} u(t,\, x,\, p)$ и первого уравнения (82) получаем

$$L_u(t,\, x_\eta(t),\, \dot{x}_\eta(t)) = L_u(t,\, x_\eta(t),\, u(t,\, x_\eta(t),\, p_\eta(t))) \overset{(80)}{=} p_\eta(t);$$

из равенства $L_x(t,\, x,\, u(t,\, x,\, p)) \overset{(81)}{=} -H_x(t,\, x,\, p)$ получаем

$$L_x(t,\, x_\eta(t),\, \dot{x}_\eta(t)) \overset{(81),(82)}{=} L_x(t,\, x_\eta(t),\, u(t,\, x_\eta(t),\, p_\eta(t))) = -H_x(t,\, x_\eta(t),\, p_\eta(t));$$

остается применить второе уравнение (82).

Итак, мы получили n-параметрическое C^2-гладкое семейство экстремалей $\{x_\eta(\cdot)\}_{\eta\in O_\sigma(p_*)}$, удовлетворяющее условию $x_\eta(t_*)=x_*$.

Наша следующая цель — доказать, что если на $[t_*,\, t_1+\gamma]$ для $\hat{x}$ выполнено усиленное условие Якоби, то при малых $\sigma>0$ построенное семейство экстремалей будет полем, покрывающим окрестность графика $\hat{x}(t)$, $t\in[t_0,\, t_1]$.

19.2 Уравнение в вариациях для системы Гамильтона и уравнение Якоби

Пусть $v\in\mathbb{R}^n$, $|\lambda|<\lambda_0$, λ_0 — достаточно малое число. Как и в предыдущем параграфе, через $(x_{p_*+\lambda v},\, p_{p_*+\lambda v})$ обозначим решение (82) с начальным условием $x(t_*)=x_*$, $p(t_*)=p_*+\lambda v$. Определим отображение $F:(-\lambda_0,\,\lambda_0)\to C([t_*,\, t_1+\delta],\, \mathbb{R}^{2n})$ равенством $F(\lambda)=(x_{p_*+\lambda v},\, p_{p_*+\lambda v})$. Из курса дифференциальных уравнений (см. также доказательство для одномерного случая при исследовании задачи оптимального управления) известно, что тогда $F'(0)=(h,\, w)=(h_1,\,\dots,\,h_n,\, w_1,\,\dots,\,w_n)$ — решение системы

$$\begin{aligned} \dot{h}_k &= \sum_{i=1}^n H_{p_kx_i}(t,\,\hat{x}(t),\,\hat{p}(t))h_i+\sum_{i=1}^n H_{p_kp_i}(t,\,\hat{x}(t),\,\hat{p}(t))w_i,\\ \dot{w}_k &= -\sum_{i=1}^n H_{x_kx_i}(t,\,\hat{x}(t),\,\hat{p}(t))h_i-\sum_{i=1}^n H_{x_kp_i}(t,\,\hat{x}(t),\,\hat{p}(t))w_i, \quad 1\le k\le n, \end{aligned} \tag{83}$$

с граничными условиями $h(t_*)=0$, $w(t_*)=v$. Покажем, что h является решением уравнения Якоби. Для этого перепишем (83) в терминах функции L и исключим переменную w.

Продифференцируем равенство $L_{u_j}(t,\, x,\, u(t,\, x,\, p))=p_j$ по p_i и по x_i:

$$\sum_{k=1}^n L_{u_ju_k}(t,\, x,\, u(t,\, x,\, p))(u_k)_{p_i}(t,\, x,\, p)=\delta_{ij},$$

$$L_{u_jx_i}(t,\, x,\, u(t,\, x,\, p))+\sum_{k=1}^n L_{u_ju_k}(t,\, x,\, u(t,\, x,\, p))(u_k)_{x_i}(t,\, x,\, p)=0.$$

Также напомним, что $u_k(t,\, x,\, p) \overset{(81)}{=} H_{p_k}(t,\, x,\, p)$. Значит,

$$\sum_{k=1}^{n} L_{u_j u_k}(t,\, x,\, u(t,\, x,\, p)) H_{p_i p_k}(t,\, x,\, p) = \delta_{ij},$$

$$L_{u_j x_i}(t,\, x,\, u(t,\, x,\, p)) + \sum_{k=1}^{n} L_{u_j u_k}(t,\, x,\, u(t,\, x,\, p)) H_{x_i p_k}(t,\, x,\, p) = 0. \qquad (84)$$

Поэтому из первого уравнения (83) и равенства $\dot{\hat{x}}(t) = u(t,\, \hat{x}(t),\, \hat{p}(t))$ получаем

$$\sum_{k=1}^{n} \hat{L}_{u_k u_j}(t) \dot{h}_k = \sum_{k=1}^{n} \sum_{i=1}^{n} \hat{L}_{u_k u_j}(t) H_{p_k x_i}(t,\, \hat{x}(t),\, \hat{p}(t)) h_i +$$

$$+ \sum_{k=1}^{n} \sum_{i=1}^{n} \hat{L}_{u_k u_j}(t) H_{p_k p_i}(t,\, \hat{x}(t),\, \hat{p}(t)) w_i = -\sum_{i=1}^{n} \hat{L}_{u_j x_i}(t) h_i + w_j,$$

т.е.

$$w_j = \sum_{i=1}^{n} \hat{L}_{u_j x_i}(t) h_i + \sum_{i=1}^{n} \hat{L}_{u_i u_j}(t) \dot{h}_i. \qquad (85)$$

Подставим это выражение во второе уравнение (83). Получаем

$$\frac{d}{dt}\left(\sum_{i=1}^{n} \hat{L}_{u_j x_i}(t) h_i + \sum_{i=1}^{n} \hat{L}_{u_i u_j}(t) \dot{h}_i \right) = \dot{w}_j \overset{(83)}{=}$$

$$= -\sum_{i=1}^{n} H_{x_j x_i}(t,\, \hat{x}(t),\, \hat{p}(t)) h_i - \sum_{k=1}^{n} H_{x_j p_k}(t,\, \hat{x}(t),\, \hat{p}(t)) w_k =$$

$$= -\sum_{i=1}^{n} H_{x_j x_i}(t,\, \hat{x}(t),\, \hat{p}(t)) h_i - \sum_{i=1}^{n} \sum_{k=1}^{n} H_{p_k x_j}(t,\, \hat{x}(t),\, \hat{p}(t)) \hat{L}_{u_k x_i}(t) h_i -$$

$$- \sum_{i=1}^{n} \sum_{k=1}^{n} H_{p_k x_j}(t,\, \hat{x}(t),\, \hat{p}(t)) \hat{L}_{u_i u_k}(t) \dot{h}_i.$$

В силу (84), $\sum\limits_{k=1}^{n} H_{p_k x_j}(t,\, \hat{x}(t),\, \hat{p}(t))\hat{L}_{u_i u_k}(t) = -\hat{L}_{u_i x_j}(t)$. Поэтому осталось доказать, что

$$\sum_{k=1}^{n} H_{p_k x_j}(t,\, \hat{x}(t),\, \hat{p}(t))\hat{L}_{u_k x_i}(t) + H_{x_j x_i}(t,\, \hat{x}(t),\, \hat{p}(t)) = -\hat{L}_{x_i x_j}(t).$$

Напомним, что $H_{x_j}(t,\, x,\, p) = -L_{x_j}(t,\, x,\, u(t,\, x,\, p))$. Сделав обратную замену переменной, получим $H_{x_j}(t,\, x,\, p(t,\, x,\, u)) = -L_{x_j}(t,\, x,\, u)$. Продифференцируем по x_i и получим

$$H_{x_j x_i}(t,\, x,\, p(t,\, x,\, u)) + \sum_{k=1}^{n} H_{x_j p_k}(t,\, x,\, p(t,\, x,\, u))(p_k)_{x_i}(t,\, x,\, u) = -L_{x_j x_i}(t,\, x,\, u).$$

Остается вспомнить, что $p_k(t,\, x,\, u) = L_{u_k}(t,\, x,\, u)$.

Замечание 1. *Из* (85) *следует, что если* $(h,\, w)$ *— решение* (83) *и* $h \equiv 0$, *то* $w \equiv 0$.

19.3 Построение центрального поля экстремалей

Ранее было доказано, что если η принадлежит достаточно малой окрестности U точки p_*, то решение $(x_\eta,\, p_\eta)$ системы (82) с начальным условием $x(t_*) = x_*$, $p(t_*) = \eta$ существует на всем отрезке $[t_*,\, t_1 + \gamma]$ и является C^2-гладким по $(t,\, \eta)$. Кроме того, было доказано, что x_η является решением уравнения Эйлера. Теперь докажем, что если окрестность U точки p_* достаточно мала, то семейство $\{x_\eta :\ \eta \in U\}$ образует поле экстремалей.

Пусть $\tau \in [t_0,\, t_1]$. Рассмотрим отображение $\Phi : (t,\, \eta) \mapsto (t,\, x_\eta(t))$, где $(t,\, \eta)$ принадлежит окрестности точки $(\tau,\, p_*)$. Если $\det \Phi'(\tau,\, p_*) \neq 0$, то, по теореме об обратном отображении, Φ на малой окрестности точки $(\tau,\, p_*)$ будет диффеоморфизмом.

Пусть $\Psi(\eta) = x_\eta(\tau)$. Из формулы для Φ следует, что достаточно проверить условие $\det \Psi'(p_*) \neq 0$. Напомним (см. предыдущий параграф), что $\Psi'_{\eta_j}(p_*)$ — это компонента $h(\tau)$ решения системы (83) с начальным условием $h(t_*) = 0$, $w(t_*) = e_j$ (j-й стандартный базисный вектор в $\mathbb{R}^n$); обозначим ее через $h(\tau,\, e_j)$. Если $\det \Psi'(p_*) = 0$, то существует ненулевой набор чисел $\lambda_1,\, \ldots,\, \lambda_n$ такой, что $\sum\limits_{j=1}^{n} \lambda_j h(\tau,\, e_j) = 0$. Значит, для решения системы (83) с начальным условием $h(t_*) = 0$, $w(t_*) = (\lambda_1,\, \ldots,\, \lambda_n)$

выполнено $h(\tau) = 0$. Если $h \equiv 0$, то в силу замечания 1 $w \equiv 0$ и $\lambda_1 = \cdots = \lambda_n = 0$. Значит, $h \neq 0$. В предыдущем параграфе было показано, что h — решение уравнения Якоби. Получаем противоречие с усиленным условием Якоби.

Теперь в каждой точке $(\tau, \hat{x}(\tau))$ возьмем окрестность V_τ, являющуюся диффеоморфным образом окрестности $W_\tau = [\tau - \varepsilon_\tau, \tau + \varepsilon_\tau] \times U_\tau$ точки (τ, p_*); здесь $\tau \in [t_0, t_1]$. Выбираем конечное подпокрытие $\{V_{\tau_k}\}_{k=1}^m$ графика $\hat{x}$ на $[t_0, t_1]$. Тогда $\{x_\eta : \ \eta \in \cap_{k=1}^m U_{\tau_k}\}$ будет искомым полем экстремалей.

19.4 Доказательство усиленного условия Якоби на расширенном отрезке

Нам осталось проверить следующий факт: если на $[t_0, t_1]$ выполнено усиленное условие Якоби, то при достаточно малых $\varepsilon > 0$ на $[t_0 - \varepsilon, t_1 + \varepsilon]$ также выполнено усиленное условие Якоби.

Как уже говорилось, если выполнено усиленное условие Лежандра, то пространство решений уравнения Якоби имеет размерность $2n$. Пусть $\varphi_{(1)}, \ldots, \varphi_{(2n)}$ — базис этого пространства. Предположим, что существует последовательность $\delta_m \downarrow 0$, $\tau_m \in (t_0 - \delta_m, t_1 + \delta_m]$, $(\alpha_1^m, \ldots, \alpha_{2n}^m) \in \mathbb{R}^{2n}$ такие, что $\|(\alpha_1^m, \ldots, \alpha_{2n}^m)\| = 1$, и функция $h^m = \sum\limits_{k=1}^{2n} \alpha_k^m \varphi_{(k)}$ удовлетворяет равенствам $h^m(t_0 - \delta_m) = h^m(\tau_m) = 0$.

Переходя к подпоследовательностям, можем считать, что $\tau_m \underset{m\to\infty}{\to} \tau$, $(\alpha_1^m, \ldots, \alpha_{2n}^m) \underset{m\to\infty}{\to} (\alpha_1, \ldots, \alpha_{2n})$. Тогда $\tau \in [t_0, t_1]$. Положим $\hat{h} = \sum\limits_{k=1}^{2n} \alpha_k \varphi_{(k)}$. Тогда $\hat{h}$ — ненулевое решение уравнения Якоби, $\hat{h}(t_0) = \hat{h}(\tau) = 0$. В силу усиленного условия Якоби, $\tau = t_0$.

Теперь мы докажем, что случай $\tau = t_0$ невозможен. Рассмотрим функционалы

$$J_m(h) = \int\limits_{t_0 - \delta_m}^{t_0 + \gamma} (\dot{h} \cdot \hat{L}_{\dot{x}\dot{x}}(t)h + 2\dot{h} \cdot \hat{L}_{\dot{x}x}(t)h(t) + h \cdot \hat{L}_{xx}(t)h)\, dt,$$

где $\gamma > 0$ — достаточно малое число. Так как $h^m(t_0 - \delta_m) = h^m(\tau_m) = 0$ и $\tau_m \underset{m\to\infty}{\to} t_0$, то при больших m функционал J_m не является неотрицательным на пространстве функций $h \in C^1([t_0 - \delta_m, t_0 + \gamma], \mathbb{R}^n)$ таких, что $h(t_0 - \delta_m) = h(t_0 + \gamma) = 0$ (см. необходимое условие Якоби и его доказательство).

Покажем, что при достаточно малом $\gamma > 0$ и больших m функционал J_m неотрицательный (и тем самым получим противоречие).

Так как $\hat{L}_{\dot{x}\dot{x}}(t_0) > 0$, то существует $a > 0$ такое, что для любого $\xi \in \mathbb{R}^n$ выполнено $\xi \cdot \hat{L}_{\dot{x}\dot{x}}(t_0)\xi \geq a\|\xi\|^2$ (здесь $\|\cdot\|$ — евклидова норма на $\mathbb{R}^n$). Значит, при малых $\gamma > 0$ и больших $m \in \mathbb{N}$ выполнено $\xi \cdot \hat{L}_{\dot{x}\dot{x}}(t)\xi \geq \frac{a}{2}\|\xi\|^2$, $t \in [t_0 - \delta_m,\, t_0 + \gamma]$.

Так как коэффициенты матриц $\hat{L}_{\dot{x}x}(t)$ и $\hat{L}_{xx}(t)$ ограничены, то в силу неравенства Коши–Буняковского существует $M > 0$ такое, что

$$|\dot{h}(t) \cdot \hat{L}_{\dot{x}x}(t)h(t)| \leq M\|\dot{h}(t)\| \cdot \|h(t)\|,$$

$$|h(t) \cdot \hat{L}_{xx}(t)h(t)| \leq M\|h(t)\|^2.$$

Далее, существует $M_1 > 0$ такое, что

$$M\|h(t)\| \cdot \|\dot{h}(t)\| \leq \frac{a}{4}\|\dot{h}(t)\|^2 + M_1\|h(t)\|^2.$$

Поэтому существует $M_2 > 0$ такое, что

$$J_m(h) \geq \int\limits_{t_0-\delta_m}^{t_0+\gamma} \left(\frac{a}{4}\sum_{j=1}^{n}\dot{h}_j^2 - M_2\sum_{j=1}^{n}h_j^2\right) dt.$$

Осталось проверить, что правая часть неотрицательна при малом $\gamma > 0$ и больших m для всех $h \in C^1([t_0 - \delta_m,\, t_0 + \gamma])$ таких, что $h(t_0 - \delta_m) = h(t_0 + \gamma) = 0$. Для этого достаточно проверить, что для любого $\omega > 0$ существует $\delta(\omega) > 0$ такое, что при $0 < \tau < \delta(\omega)$

$$\int\limits_0^{\tau} (\dot{x}^2 - \omega^2 x^2)\, dt \geq 0, \quad x \in C^1[0,\, \tau], \quad x(0) = x(\tau) = 0. \qquad (86)$$

Это уже было доказано в §1, когда разбирался пример с гармоническим осциллятором.

20 Доказательство неравенств с помощью проверки формулы Вейерштрасса

Рассмотрим пример. Нужно доказать, что для любой функции $x \in C^1[0,\, +\infty)$ такой, что $x(0) = 0$ и $x(t) = o(\sqrt{t})$ при $t \to +\infty$, выполнено неравенство

Гильберта:

$$\int_0^\infty \dot{x}^2\,dt \geq \int_0^\infty \frac{x^2}{4t^2}\,dt.$$

Решать эту задачу можно следующим образом. Рассмотрим функционал

$$\mathcal{L}(x) = \int_0^\infty \left(\dot{x}^2 - \frac{x^2}{4t^2}\right)\,dt$$

и найдем для него однопараметрическое семейство экстремалей $x(t,\,\alpha)$ таких, что $x(0,\,\alpha) = 0$ для любого $\alpha \in \mathbb{R}$. Для $\tau > 0$, $\xi > 0$ найдем функцию наклона поля $u(\tau,\,\xi)$. Затем с помощью интегрирования по частям проверим, что выполнена формула Вейерштрасса. (Здесь непосредственно применить теорию мы не можем, так как здесь бесконечный интервал и функция L имеет особенность.)

Уравнение Эйлера–Лагранжа имеет вид $-\frac{d}{dt}2\dot{x} - \frac{x}{2t^2} = 0$, т.е. $\ddot{x} + \frac{x}{4t^2} = 0$. Подберем решение в виде $x(t) = t^\beta$. Подставив в уравнение, получим $\beta = \frac{1}{2}$. Получили семейство экстремалей $x(t,\,\alpha) = \alpha t^{1/2}$.

Найдем функцию наклона поля. Если $\alpha(\tau,\,\xi)\tau^{1/2} = \xi$, то $\alpha(\tau,\,\xi) = \frac{\xi}{\tau^{1/2}}$. Значит, $u(\tau,\,\xi) = \frac{\xi}{\tau^{1/2}} \cdot \frac{1}{2}\tau^{-1/2} = \frac{\xi}{2\tau}$.

Найдем функцию Вейерштрасса: $\mathcal{E}(t,\,x,\,u,\,v) = v^2 - u^2 - 2u(v-u) = (v-u)^2$.

Таким образом, нам нужно проверить, что

$$\int_0^\infty \left(\dot{x}^2 - \frac{x^2}{4t^2}\right)\,dt = \int_0^\infty \left(\dot{x} - \frac{x}{2t}\right)^2\,dt,$$

где $x \in C^1[0,\,+\infty)$, $x(0) = 0$, $x(t) = o(\sqrt{t})$ при $t \to +\infty$. Если это равенство верно, то левая часть неотрицательна и требуемое утверждение будет доказано.

Раскроем скобки в правой части и проинтегрируем по частям:

$$\int_0^\infty \left(\dot{x} - \frac{x}{2t}\right)^2\,dt = \int_0^\infty \left(\dot{x}^2 - \frac{1}{2t}\cdot 2x\dot{x} + \frac{x^2}{4t^2}\right)\,dt =$$

$$= \int_0^\infty \left(\dot{x}^2 + \frac{x^2}{4t^2}\right)\,dt - \int_0^\infty \frac{x^2}{2t^2}\,dt + \left.\frac{x^2}{2t}\right|_0^\infty = \int_0^\infty \left(\dot{x}^2 - \frac{x^2}{4t^2}\right)\,dt + \left.\frac{x^2}{2t}\right|_0^\infty.$$

Остается проверить, что $\left.\frac{x^2}{2t}\right|_0^\infty = 0$. Так как $x(0) = 0$ и $x \in C^1[0, \infty)$, то $x(t) = O(t)$ при $t \to 0$; значит, $\frac{x^2(t)}{t} = \frac{O(t^2)}{t} = O(t) \underset{t\to 0}{\to} 0$. Далее, нам дано, что $x(t) = o(\sqrt{t})$, $t \to +\infty$. Значит, $\frac{x^2(t)}{t} = \frac{o(t)}{t} = o(1)$ при $t \to +\infty$.

21 Теоремы существования точки минимума

21.1 Теорема Вейерштрасса, рефлексивные пространства и слабая сходимость

Пусть T — топологическое пространство, $f : T \to \mathbb{R} \cup \{+\infty\}$. Функция f называется секвенциально полунепрерывной снизу, если для любого $x \in T$, для любой последовательности $x_n \underset{n\to\infty}{\to} x$ выполнено $f(x) \le \underline{\lim}_{n\to\infty} f(x_n)$.

Скажем, что T секвенциально компактно, если для любой последовательности $\{x_n\}_{n\in\mathbb{N}} \subset T$ существует сходящаяся подпоследовательность.

Предложение 27. *(теорема Вейерштрасса). Пусть T секвенциально компактно, $f : T \to \mathbb{R} \cup \{+\infty\}$ секвенциально полунепрерывна снизу. Тогда существует точка $\hat{x} \in T$ такая, что $f(\hat{x}) = \min_{x\in T} f(x)$.*

Доказательство. Пусть $f(x_n) \underset{n\to\infty}{\to} \inf_{x\in T} f(x)$. Так как T секвенциально компактно, то существует сходящаяся подпоследовательность $\{x_{n_k}\}_{k\in\mathbb{N}}$. Положим $\hat{x} = \lim\limits_{k\to\infty} x_{n_k}$. Так как f секвенциально полунепрерывна снизу, то $f(\hat{x}) \le \lim\limits_{k\to\infty} f(x_{n_k}) = \inf_{x\in T} f(x)$. Значит, $f(\hat{x}) = \inf_{x\in T} f(x)$. □

Пространства функций бесконечномерные, поэтому шары в них некомпактны относительно топологии, заданной нормой. Но в некоторых случаях замкнутые шары слабо секвенциально компактны.

Определение 12. *Пусть X — нормированное пространство. Рассмотрим оператор изометрического вложения во второе сопряженное $\mathbf{i} : X \to X^{**}$, заданный по формуле $\mathbf{i}(x)(f) = f(x)$, $f \in X^*$. Пространство X называется рефлексивным, если $\mathbf{i}$ сюръективен.*

Пример. Гильбертово пространство рефлексивно.

Пример. Пусть $X = L_p(\Omega,\,\mu)$, $1 < p < \infty$. Из функционального анализа известно, что $X^* = L_{p'}(\Omega,\,\mu)$, где $\frac{1}{p}+\frac{1}{p'}=1$; а именно, любой линейный непрерывный функционал на $L_p(\Omega,\,\mu)$ имеет вид $f(x)=\int\limits_\Omega x\cdot y\,d\mu$, где $y \in L_{p'}(\Omega,\,\mu)$, при этом $\|f\|_{X^*} = \|y\|_{L_{p'}(\Omega,\,\mu)}$. Если теперь применить ту же теорему к $L_{p'}(\Omega,\,\mu)$, получим, что X рефлексивно.

Напомним, что слабая топология на нормированном пространстве X задается базой окрестностей

$$U_{\varepsilon,f_1,\,\dots,\,f_n}(x)=\{y\in X:\ |f_j(x-y)|<\varepsilon,\ 1\le j\le n\},$$

где $\varepsilon>0$, $n\in\mathbb{N}$, $f_1,\,\dots,\,f_n\in X^*$. Слабая сходимость — это сходимость в слабой топологии; это эквивалентно следующему условию (часто его берут в качестве исходного определения слабой сходимости): $x_n \underset{n\to\infty}{\overset{w}{\to}} x$ тогда и только тогда, когда для любого $f\in X^*$ выполнено $f(x_n)\underset{n\to\infty}{\to} f(x)$.

Факт из функционального анализа: слабо сходящаяся последовательность ограничена.

(Это утверждение выводилось из теоремы Банаха–Штейнгауза.)

Следующую теорему приведем без доказательства.

Теорема 26. *Пусть X — банахово пространство. Следующие условия эквивалентны:*

1. *X рефлексивно;*
2. *единичный шар B_X компактен в слабой топологии;*
3. *единичный шар B_X секвенциально компактен в слабой топологии.*

(Напомним, что компактность определяется в терминах открытых покрытий.)

Импликация $1 \Rightarrow 2$ является следствием теоремы Банаха–Алаоглу о $*$-слабой компактности единичного шара в сопряженном пространстве. Эквивалентность $2 \Leftrightarrow 3$ — это теорема Эберлейна–Шмульяна. Отметим, что для метрических пространств компактность эквивалентна секвенциальной компактности. Но на бесконечномерном пространстве слабая топология не метризуема, поэтому теорему Эберлейна–Шмульяна нельзя непосредственно вывести из теоремы для метрических пространств. Для $*$-слабой топологии компактность уже не эквивалентна секвенциальной компактности.

Следствие 2. *Пусть X — рефлексивное банахово пространство, $Y \subset X$ — замкнутое подпространство. Тогда Y рефлексивно.*

Доказательство. Достаточно показать, что единичный шар в Y секвенциально слабо компактен. Пусть $\{x_n\}_{n\in\mathbb{N}} \subset B_Y = B_X \cap Y$. Так как X рефлексивно, то существует подпоследовательность $x_{n_k} \underset{k\to\infty}{\overset{w}{\to}} x \in B_X$ в X. Докажем, что $x \in Y$. В самом деле, иначе по теореме отделимости существует функционал $x^* \in X^*$ такой, что $x^*(x) > \sup_{y\in Y} x^*(y) \geq \lim_{k\to\infty} x^*(x_{n_k})$ — противоречие. $\square$

Из теоремы Вейерштрасса и теоремы 26 получается следующее утверждение.

Теорема 27. *Пусть X — рефлексивное банахово пространство.*

1. *Пусть $B \subset X$ — замкнутый шар, $f : B \to \mathbb{R}\cup\{+\infty\}$ секвенциально слабо полунепрерывна снизу. Тогда f имеет точку минимума на B.*

2. *Пусть $f : X \to \mathbb{R} \cup \{+\infty\}$ секвенциально слабо полунепрерывна снизу, при этом $f(x) \underset{\|x\|\to+\infty}{\to} +\infty$. Тогда f имеет точку минимума на X.*

Пункт 2 следует из пункта 1, так как существует замкнутый шар $B \subset X$ такой, что $\inf_B f = \inf_X f$.

Условие $f(x) \underset{\|x\|\to+\infty}{\to} +\infty$ называется условием коэрцитивности.

21.2 Теорема Мазура. Достаточные условия слабой секвенциальной полунепрерывности снизу.

Пусть X — нормированное пространство, $C \subset X$. Выпуклой оболочкой множества C называется

$$\operatorname{conv} C = \left\{ \sum_{j=1}^{n} \lambda_j x_j : \; n \in \mathbb{N}, \, x_j \in C, \, \lambda_j \geq 0, \; 1 \leq j \leq n, \, \sum_{j=1}^{n} \lambda_j = 1 \right\}.$$

Тогда $\operatorname{conv} C$ — выпуклое множество. Через $\overline{\operatorname{conv}}\, C$ обозначим замыкание выпуклой оболочки C; это выпуклое замкнутое множество.

Теорема 28. *(Мазур). Пусть* $x_n \underset{n\to\infty}{\overset{w}{\to}} x$. *Тогда существует последовательность* $\{y_k\}_{k\in\mathbb{N}} \subset \operatorname{conv}\{x_n\}_{n\in\mathbb{N}}$ *такая, что* $\|y_k - x\| \underset{k\to\infty}{\to} 0$.

Доказательство. Пусть $M = \overline{\operatorname{conv}}\{x_n\}_{n\in\mathbb{N}}$. Если $x \in M$, то утверждение сразу следует из определения замыкания.

Пусть $x \notin M$. По теореме отделимости, существует функционал $x^* \in X^*$ такой, что $x^*(x) > \sup_{y\in M} x^*(y)$; в частности, $x^*(x) > \sup_{n\in\mathbb{N}} x^*(x_n)$. Но $x_n \underset{n\to\infty}{\overset{w}{\to}} x$, так что $x^*(x_n) \underset{n\to\infty}{\to} x^*(x)$. Противоречие. □

Отметим следующее свойство выпуклых функций: если $f : X \to \mathbb{R} \cup \{+\infty\}$, $x_1, \ldots, x_n \in X$, $\lambda_1, \ldots, \lambda_n \geq 0$, $\sum\limits_{j=1}^{n} \lambda_j = 1$, то

$$f\left(\sum_{j=1}^{n} \lambda_j x_j\right) \leq \sum_{j=1}^{n} \lambda_j f(x_j);$$

это доказывается индукцией по n.

Теорема 29. *Пусть функция* $f : X \to \mathbb{R} \cup \{+\infty\}$ *выпуклая и секвенциально полунепрерывная снизу относительно нормы. Тогда* f *слабо секвенциально полунепрерывна снизу.*

Доказательство. Пусть $x_n \underset{n\to\infty}{\overset{w}{\to}} x$, $f(x_n) \underset{n\to\infty}{\to} a$. Покажем, что $f(x) \leq a$.

Для любого $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что для любого $n \geq N$ выполнено $f(x_n) \leq a + \varepsilon$. По теореме Мазура, существует последовательность $\{y_k\}_{k\in\mathbb{N}} \subset \operatorname{conv}\{x_n\}_{n\geq N}$ такая, что $\|y_k - x\| \underset{k\to\infty}{\to} 0$. Покажем, что $f(y_k) \leq a + \varepsilon$, $k \in \mathbb{N}$ (тогда $f(x) \leq a + \varepsilon$ в силу секвенциальной полунепрерывности снизу; в силу произвольности ε, получаем $f(x) \leq a$).

Имеем $y_k = \sum\limits_{j=N}^{N+m} \lambda_j x_j$, $\lambda_j \geq 0$, $\sum\limits_{j=N}^{N+m} \lambda_j = 1$. Так как f выпукла, то $f(y_k) \leq \sum\limits_{j=N}^{N+m} \lambda_j f(x_j) \leq a + \varepsilon$. □

Отсюда получаем следующее утверждение.

Теорема 30. *Пусть* X *— рефлексивное банахово пространство,* $f : X \to \mathbb{R} \cup \{+\infty\}$ *— выпуклая секвенциально полунепрерывная снизу относительно нормы функция. Пусть* $f(x) \to +\infty$ *при* $\|x\| \to +\infty$. *Тогда* f *имеет точку минимума на* X.

Пример. Пусть X — рефлексивное банахово пространство, $C \subset X$ — непустое выпуклое замкнутое множество, $x_0 \in X$. Положим

$$f(x) = \begin{cases} \|x - x_0\|, & x \in C, \\ +\infty, & x \notin C. \end{cases}$$

Тогда f — выпуклая секвенциально полунепрерывная снизу функция (здесь используется выпуклость и замкнутость C), при этом f коэрцитивна. Значит, f имеет точку минимума $\hat{x} \in C$. Получаем $\|\hat{x} - x_0\| = \min_{x \in X} f(x) = \min_{x \in C} \|x - x_0\|$, т.е. $\hat{x}$ — ближайшая точка к x_0 из C. Таким образом, в рефлексивном пространстве любая точка имеет ближайшую из выпуклого замкнутого непустого множества.

21.3 Пространства Соболева

Мы будем для простоты рассматривать пространства Соболева с нулевыми граничными условиями.

Пусть $1 \le p < \infty$. Пространством Соболева $\mathring{W}_p^1([t_0, t_1], \mathbb{R}^n)$ называется пространство

$$\{x \in AC([t_0, t_1], \mathbb{R}^n) : \ \dot{x} \in L_p([t_0, t_1], \mathbb{R}^n), \ x(t_0) = x(t_1) = 0\}$$

с нормой

$$\|x\|_{\mathring{W}_p^1([t_0, t_1], \mathbb{R}^n)} = \left(\int\limits_{t_0}^{t_1} \sum_{j=1}^{n} |\dot{x}_j(t)|^p \, dt \right)^{1/p}.$$

Определим также норму на $L_p([t_0, t_1], \mathbb{R}^n)$ по формуле

$$\|x\|_{L_p([t_0, t_1], \mathbb{R}^n)} = \left(\int\limits_{t_0}^{t_1} \sum_{j=1}^{n} |x_j(t)|^p \, dt \right)^{1/p}.$$

Пространство $\mathring{W}_p^1([t_0, t_1], \mathbb{R}^n)$ изометрически изоморфно пространству

$$L = \left\{ y \in L_p([t_0, t_1], \mathbb{R}^n) : \ \int\limits_{t_0}^{t_1} y_i(t) \, dt = 0, \ 1 \le i \le n \right\};$$

изоморфизм каждой функции $x \in \mathring{W}_p^1([t_0, t_1], \mathbb{R}^n)$ сопоставляет функцию $\dot{x}$. Если $1 < p < \infty$, то $L_p([0, 1], \mathbb{R}^n)$ рефлексивно. Так как L — замкнутое подпространство в $L_p([0, 1], \mathbb{R}^n)$, то L тоже рефлексивно (по следствию 2). Мы получили

Предложение 28. *Пусть $1 < p < \infty$. Тогда $\mathring{W}_p^1([t_0, t_1], \mathbb{R}^n)$ рефлексивно.*

Следствие 3. *Пусть $1 < p < \infty$, $f : \mathring{W}_p^1([t_0, t_1], \mathbb{R}^n) \to \mathbb{R} \cup \{+\infty\}$ слабо секвенциально полунепрерывна снизу и коэрцитивна. Тогда f имеет точку минимума.*

Следствие 4. *Пусть $1 < p < \infty$, $f : \mathring{W}_p^1([t_0, t_1], \mathbb{R}^n) \to \mathbb{R} \cup \{+\infty\}$ секвенциально полунепрерывна снизу относительно нормы, выпукла и коэрцитивна. Тогда f имеет точку минимума.*

Из определения следует, что $\mathring{W}_p^1([t_0, t_1], \mathbb{R}^n)$ содержится в $C([t_0, t_1], \mathbb{R}^n)$ как множество. Пусть $J : \mathring{W}_p^1([t_0, t_1], \mathbb{R}^n) \to C([t_0, t_1], \mathbb{R}^n)$ — оператор вложения, то есть $J(x) = x$.

Предложение 29. *Пусть $1 < p < \infty$. Тогда оператор J компактен.*

Доказательство. Нужно показать, что множество

$$B = \{x \in \mathring{W}_p^1([t_0, t_1], \mathbb{R}^n) : \ \|x\|_{\mathring{W}_p^1([t_0, t_1], \mathbb{R}^n)} \le 1\}$$

предкомпактно в $C([t_0, t_1], \mathbb{R}^n)$. По теореме Арцела–Асколи, достаточно проверить, что это множество равномерно ограничено и равностепенно непрерывно.

Для $k \in \{1, \ldots, n\}$, $t_0 \le t < \tau \le t_1$, $x \in B$ имеем

$$|x_k(\tau) - x_k(t)| = \left| \int\limits_t^\tau \dot{x}_k(s)\, ds \right| \le (\tau - t)^{1/p'} \|\dot{x}_k\|_{L_p[t, \tau]} \le (\tau - t)^{1/p'},$$

откуда следует равностепенная непрерывность. Подставив $t = t_0$ и учитывая, что $x(t_0) = 0$, из этой же оценки получаем равномерную ограниченность. □

Напомним следующий факт из функционального анализа: если оператор компактен, то он переводит слабо сходящуюся последовательность в сходящуюся по норме.

Следствие 5. *Если* $x_k \underset{k\to\infty}{\overset{w}{\to}} x$ *в* $\mathring{W}_p^1([t_0,\, t_1],\, \mathbb{R}^n)$, *то* x_k *равномерно сходится к* x.

Нам в дальнейшем также понадобится

Предложение 30. *Пусть* $x \in L_{p'}([t_0,\, t_1],\, \mathbb{R}^n)$, $f(h) = \int\limits_{t_0}^{t_1} x(t)\dot{h}(t)\, dt$, $h \in \mathring{W}_p^1([t_0,\, t_1],\, \mathbb{R}^n)$. *Тогда* $f \in (\mathring{W}_p^1([t_0,\, t_1],\, \mathbb{R}^n))^*$ *и*

$$|f(h)| \le \|x\|_{L_{p'}([t_0,\, t_1],\, \mathbb{R}^n)} \|h\|_{\mathring{W}_p^1([t_0,\, t_1],\, \mathbb{R}^n)}. \tag{87}$$

Это утверждение следует из неравенства Гёльдера и определения нормы в $\mathring{W}_p^1([t_0,\, t_1],\, \mathbb{R}^n)$.

22 Существование точки минимума в задаче Дидоны

Сначала рассмотрим задачу:

$$\begin{cases} f_0(x,\, y) := \int\limits_0^1 (x\dot{y} - y\dot{x})\, dt \to \inf, \\ f_1(x,\, y) := \int\limits_0^1 (\dot{x}^2 + \dot{y}^2)\, dt \le l^2, \\ (x,\, y) \in \mathring{W}_2^1([0,\, 1],\, \mathbb{R}^2). \end{cases} \tag{88}$$

В силу теоремы 27, достаточно проверить, что функция f_0 секвенциально слабо полунепрерывна снизу.

В самом деле, пусть $(x_n,\, y_n) \underset{n\to\infty}{\overset{w}{\to}} (x,\, y)$. Тогда $(x_n,\, y_n) \underset{n\to\infty}{\to} (x,\, y)$ в $C([0,\, 1],\, \mathbb{R}^2)$ (см. следствие 5). Кроме того, $\{(x_n,\, y_n)\}_{n\in\mathbb{N}}$ ограничена (т.к. это слабо сходящаяся последовательность), то есть существует $R > 0$ такое, что

$$\|(x_n,\, y_n)\|_{\mathring{W}_2^1([0,\,1],\, \mathbb{R}^2)} \le R, \quad n \in \mathbb{N}. \tag{89}$$

Получаем:

$$\left| \int\limits_0^1 (x_n\dot{y}_n - y_n\dot{x}_n)\, dt - \int\limits_0^1 (x\dot{y} - y\dot{x})\, dt \right| \le$$

$$\leq \left|\int_0^1 (x_n - x)\dot{y}_n\,dt\right| + \left|\int_0^1 x(\dot{y}_n - \dot{y})\,dt\right| + \left|\int_0^1 (y_n - y)\dot{x}_n\,dt\right| + \left|\int_0^1 y(\dot{x}_n - \dot{x})\,dt\right| \leq$$

$$\leq \|x_n - x\|_{L_2[0,\,1]}\|\dot{y}_n\|_{L_2[0,\,1]} + \|y_n - y\|_{L_2[0,\,1]}\|\dot{x}_n\|_{L_2[0,\,1]} +$$

$$+ \left|\int_0^1 x(\dot{y}_n - \dot{y})\,dt\right| + \left|\int_0^1 y(\dot{x}_n - \dot{x})\,dt\right| \overset{(89)}{\leq}$$

$$\leq R(\|x_n - x\|_{L_2[0,\,1]} + \|y_n - y\|_{L_2[0,\,1]}) + \left|\int_0^1 x(\dot{y}_n - \dot{y})\,dt\right| + \left|\int_0^1 y(\dot{x}_n - \dot{x})\,dt\right| \underset{n\to\infty}{\to} 0$$

(см. определение слабой сходимости и предложение 30). Таким образом,

$$\int_0^1 (x_n\dot{y}_n - y_n\dot{x}_n)\,dt \underset{n\to\infty}{\to} \int_0^1 (x\dot{y} - y\dot{x})\,dt,$$

поэтому f_0 слабо секвенциально полунепрерывен снизу.

В силу теоремы 27, точка минимума в задаче (88) существует. Найдем ее.

Вычислив вариацию по Лагранжу у f_0, f_1, получаем:

$$\begin{gathered} f_0'(x,\,y)[(h,\,w)] = \int_0^1 (h\dot{y} + x\dot{w} - w\dot{x} - y\dot{h})\,dt, \\ f_1'(x,\,y)[(h,\,w)] = \int_0^1 (2\dot{x}\dot{h} + 2\dot{y}\dot{w})\,dt. \end{gathered} \tag{90}$$

Отсюда и из (87) следует, что f_0 и f_1 непрерывно дифференцируемы. Значит, можно применить принцип Лагранжа и получить: если $(x,\,y)$ — точка минимума, то существуют $(\lambda_0,\,\lambda_1) \neq (0,\,0)$ такие, что $\lambda_0 \geq 0$, $\lambda_1 \geq 0$, $\lambda_1(f_1(x,\,y) - l^2) = 0$,

$$\lambda_0 f_0'(x,\,y)[(h,\,w)] + \lambda_1 f_1'(x,\,y)[(h,\,w)] = 0 \tag{91}$$

для любого $(h,\,w) \in \mathring{W}_2^1([0,\,1],\,\mathbb{R}^2)$. Подставив (90) в (91) и проинтегрировав по частям $\int_0^1 h\dot{y}\,dt$ и $\int_0^1 w\dot{x}\,dt$ и учитывая, что $h(0) = h(1) = w(0) =$

$w(1) = 0$, получаем

$$\int\limits_0^1 (-2\lambda_0 y\dot{h} + 2\lambda_0 x\dot{w} + 2\lambda_1 \dot{x}\dot{h} + 2\lambda_1 \dot{y}\dot{w})\, dt = 0.$$

В силу обобщенной леммы Дюбуа-Реймона, $-\lambda_0 y + \lambda_1 \dot{x} = c_1$, $\lambda_0 x + \lambda_1 \dot{y} = c_2$.

Если $\lambda_1 = 0$, то x и y будут константами, и $f_0(x,\, y) = 0$ — не минимальное значение (можно взять параметризацию простой замкнутой гладкой кривой с ориентацией против часовой стрелки; тогда $-f_0(x,\, y)/2$ — это площадь области, ограниченной этой кривой).

Значит, $\lambda_1 > 0$. Можно считать, что $\lambda_1 = 1$. Получаем систему: $\dot{x} = \lambda_0 y + c_1$, $\dot{y} = -\lambda_0 x + c_2$. Если $\lambda_0 = 0$, то опять в силу граничных условий получаем $x = 0$, $y = 0$ — это не точка минимума. Значит, $\lambda_0 > 0$. Тогда $x = a\cos\lambda_0 t + b\sin\lambda_0 t + \gamma_1$, $y = -a\sin\lambda_0 t + b\cos\lambda_0 t + \gamma_2$. Из граничных условий получаем, что $\lambda_0 = 2\pi n$, $n \in \mathbb{N}$. В силу условия дополняющей нежесткости, $f_1(x,\, y) = l^2$. Вычислив значения f_0 и f_1, получим, что минимум будет при $n = 1$.

Заметим, что найденная точка минимума является C^1-гладкой и $\dot{x}^2 + \dot{y}^2 \equiv l^2$. Значит, в задаче (88) и в задаче

$$\begin{gathered} \int_0^1 (x\dot{y} - y\dot{x})\, dt \to \inf, \quad \dot{x}^2 + \dot{y}^2 = l^2, \\ x(0) = x(1) = y(0) = y(1) = 0, \quad (x,\, y) \in C^1([0,\, 1],\, \mathbb{R}^2) \end{gathered} \tag{92}$$

точка минимума одна и та же. Теперь рассмотрим задачу Дидоны:

$$\int\limits_0^1 (x\dot{y} - y\dot{x})\, dt \to \inf, \quad \int\limits_0^1 \sqrt{\dot{x}^2 + \dot{y}^2}\, dt = l,$$

$$x(0) = x(1) = y(0) = y(1) = 0, \quad (x,\, y) \in C^1([0,\, 1],\, \mathbb{R}^2), \quad \dot{x}^2 + \dot{y}^2 > 0.$$

В ней оба функционала инвариантны относительно изменения параметризации. Значит, достаточно искать минимум на множестве кривых с параметризацией, пропорциональной натуральной, т.е. таких, что $\dot{x}^2 + \dot{y}^2 = l^2$. Поэтому задача Дидоны эквивалентна задаче (92). Следовательно, в ней точка минимума существует и совпадает с точкой минимума в (88), т.е. является окружностью.

23 Теорема Тонелли

Пусть $1 < p < \infty$. Рассмотрим задачу

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt \to \inf, \quad x \in \mathring{W}_p^1[t_0,\, t_1]. \tag{93}$$

Теорема 31. *Пусть $L \in C([t_0,\, t_1] \times \mathbb{R} \times \mathbb{R})$ выпукла по $\dot{x}$, $L_{\dot{x}} \in C([t_0,\, t_1] \times \mathbb{R} \times \mathbb{R})$, существуют $C > 0$ и $\beta > 0$ такие, что*

$$L(t,\, \xi,\, \eta) \geq C|\eta|^p - \beta. \tag{94}$$

Тогда точка минимума в (93) существует.

Доказательство. В силу (94),

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt \geq C \int\limits_{t_0}^{t_1} |\dot{x}(t)|^p\, dt - \beta(t_1 - t_0) =$$

$$= C\|x\|^p_{\mathring{W}_p^1[t_0,\, t_1]} - \beta(t_1 - t_0) \underset{\|x\|_{\mathring{W}_p^1[t_0,\, t_1]} \to \infty}{\longrightarrow} +\infty.$$

Поэтому (см. теорему 27) достаточно доказать, что функционал $\mathcal{L}(x) = \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt$ слабо секвенциально полунепрерывен снизу.

Пусть $x_n \underset{n \to \infty}{\overset{w}{\to}} x$ в $\mathring{W}_p^1[t_0,\, t_1]$. Нужно проверить, что

$$\varliminf_{n \to \infty} \int\limits_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \geq \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt. \tag{95}$$

Переходя к подпоследовательностям, можно считать, что существует $\lim\limits_{n \to \infty} \int\limits_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt < \infty$.

В силу следствия 5,

$$\|x_n - x\|_{C[t_0,\, t_1]} \underset{n \to \infty}{\longrightarrow} 0. \tag{96}$$

В частности, $|x_n(t)|$ ограничено на $[t_0,\, t_1]$ равномерно по n и t.

При доказательстве ключевым будет неравенство

$$L(t,\, \xi,\, \eta) - L(t,\, \xi,\, \eta_0) \geq L_{\dot{x}}(t,\, \xi,\, \eta_0)(\eta - \eta_0), \tag{97}$$

которое следует из выпуклости L по η.

Для каждого $\varepsilon > 0$ положим

$$E_\varepsilon = \{t \in [t_0,\, t_1] : \ |\dot{x}(t)| \leq 1/\varepsilon\}.$$

Тогда $\gamma_\varepsilon := \mu([t_0,\, t_1] \backslash E_\varepsilon) \underset{\varepsilon \to 0+}{\to} 0$.

Пусть для каждого $\varepsilon > 0$ доказано неравенство

$$\varliminf_{n \to \infty} \int\limits_{E_\varepsilon} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \geq \int\limits_{E_\varepsilon} L(t,\, x(t),\, \dot{x}(t))\, dt. \tag{98}$$

Выведем отсюда (95). Сначала заметим, что

$$\int\limits_{[t_0,\, t_1] \backslash E_\varepsilon} L(t,\, x_n,\, \dot{x}_n)\, dt \overset{(94)}{\geq} -\beta\gamma_\varepsilon \underset{\varepsilon \to 0}{\to} 0.$$

Для каждого $\alpha > 0$ выберем $\varepsilon_\alpha > 0$ такое, что при $\varepsilon \in (0,\, \varepsilon_\alpha]$ выполнено $\beta\gamma_\varepsilon < \alpha$. Для $\varepsilon > 0$ в соответствии с (98) выбираем $N(\varepsilon) \in \mathbb{N}$ такое, что при $n \geq N(\varepsilon)$

$$\int\limits_{E_\varepsilon} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \geq \int\limits_{E_\varepsilon} L(t,\, x(t),\, \dot{x}(t))\, dt - \alpha.$$

Тогда при $\varepsilon \in (0,\, \varepsilon_\alpha]$, $n \geq N(\varepsilon)$

$$\int\limits_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \geq \int\limits_{E_\varepsilon} L(t,\, x(t),\, \dot{x}(t))\, dt - 2\alpha.$$

Значит, при $\varepsilon \in (0,\, \varepsilon_\alpha]$

$$\varliminf_{n \to \infty} \int\limits_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \geq \int\limits_{E_\varepsilon} L(t,\, x(t),\, \dot{x}(t))\, dt - 2\alpha.$$

Отсюда следует, что

$$\varliminf_{n\to\infty}\int\limits_{t_0}^{t_1} L(t,\, x_n(t),\, \dot x_n(t))\, dt \ge \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot x(t))\, dt - 2\alpha.$$

В силу произвольности α, получаем (95).

Теперь докажем (98). Имеем

$$\int\limits_{E_\varepsilon} (L(t,\, x_n,\, \dot x_n) - L(t,\, x,\, \dot x))\, dt =$$

$$= \int\limits_{E_\varepsilon} (L(t,\, x_n,\, \dot x_n) - L(t,\, x_n,\, \dot x))\, dt + \int\limits_{E_\varepsilon} (L(t,\, x_n,\, \dot x) - L(t,\, x,\, \dot x))\, dt.$$

Оценим первое слагаемое:

$$\int\limits_{E_\varepsilon} (L(t,\, x_n,\, \dot x_n) - L(t,\, x_n,\, \dot x))\, dt \overset{(97)}{\ge} \int\limits_{E_\varepsilon} L_{\dot x}(t,\, x_n,\, \dot x)(\dot x_n - \dot x)\, dt =$$

$$= \int\limits_{E_\varepsilon} L_{\dot x}(t,\, x,\, \dot x)(\dot x_n - \dot x)\, dt + \int\limits_{E_\varepsilon} (L_{\dot x}(t,\, x_n,\, \dot x) - L_{\dot x}(t,\, x,\, \dot x))(\dot x_n - \dot x)\, dt.$$

Напомним, что x_n слабо сходится к x в $\mathring{W}_p^1[t_0,\, t_1]$. Функция $L_{\dot x}(t,\, x(t),\, \dot x(t))$ ограничена на E_ε. Значит, в силу предложения 30

$$\int\limits_{E_\varepsilon} L_{\dot x}(t,\, x(t),\, \dot x(t))(\dot x_n(t) - \dot x(t))\, dt \underset{n\to\infty}{\to} 0.$$

Последовательность $\dot x_n - \dot x$ ограничена в $L_p[t_0,\, t_1]$; из (96) следует, что $L_{\dot x}(t,\, x_n(t),\, \dot x(t)) - L_{\dot x}(t,\, x(t),\, \dot x(t))$ равномерно на E_ε сходится к 0 (так как $L_{\dot x}$ непрерывна, то она равномерно непрерывна на компактах). Значит,

$$\int\limits_{E_\varepsilon} (L_{\dot x}(t,\, x_n,\, \dot x) - L_{\dot x}(t,\, x,\, \dot x))(\dot x_n - \dot x)\, dt \underset{n\to\infty}{\to} 0.$$

Таким образом,

$$\varliminf_{n\to\infty}\int\limits_{E_\varepsilon} (L(t,\, x_n,\, \dot x_n) - L(t,\, x_n,\, \dot x))\, dt \ge 0.$$

Теперь рассмотрим второе слагаемое:

$$\int\limits_{E_\varepsilon} (L(t,\, x_n,\, \dot{x}) - L(t,\, x,\, \dot{x}))\, dt \underset{n\to\infty}{\to} 0,$$

так как $L(t,\, x_n,\, \dot{x}) - L(t,\, x,\, \dot{x})$ равномерно сходится к 0 на E_ε (так как L непрерывна, то она равномерно непрерывна на компактах).

В итоге получаем (98). $\square$

Пример. Рассмотрим задачу

$$\mathcal{L}(x) := \int\limits_0^1 ((1-\dot{x}^2)^2 + x^2)\, dt \to \inf, \quad x \in \mathring{W}_4^1[0,\, 1].$$

Тогда $\mathcal{L}(x) > 0$ для любой функции $x \in \mathring{W}_4^1[0,\, 1]$ и

$$\inf\{\mathcal{L}(x) :\ x \in \mathring{W}_4^1[0,\, 1]\} = 0$$

(это доказывается точно так же для липшицевых функций, см. задачу 2 п. 2). То есть $\mathcal{L}$ не имеет точки минимума на $\mathring{W}_4^1[0,\, 1]$; здесь нарушилось условие выпуклости по $\dot{x}$.