1 Дифференцирование в нормированных пространствах

В этом курсе мы рассматриваем линейные пространства только над полем $\mathbb{R}$.

Напоминание.

Нормой на линейном пространстве X называется функция $\|\cdot\| : X \to \mathbb{R}_+$ со следующими свойствами:

1. $\|x\| = 0$ тогда и только тогда, когда $x = 0$;
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$, $\lambda \in \mathbb{R}$, $x \in X$;
3. $\|x + y\| \leq \|x\| + \|y\|$, $x, y \in X$.

Норма на X порождает метрику: $\rho(x, y) = \|x - y\|$.

Нормированное пространство называется банаховым, если оно полно.

Пусть X, Y линейные нормированные пространства. Отображение $A : X \to Y$ называется линейным, если $A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$, $x, y \in X$, $\alpha, \beta \in \mathbb{R}$. Линейное отображение непрерывно тогда и только тогда, когда оно ограничено, т.е. $\sup_{\|x\|_X \leq 1} \|Ax\|_Y < \infty$.

Обозначим через $L(X, Y)$ пространство линейных непрерывных операторов из X в Y, снабженное нормой

$$\|A\| = \sup_{\|x\|_X \leq 1} \|Ax\|_Y.$$

Сопряженным к нормированному пространству X (обозначение: X^*) называется пространство линейных непрерывных функционалов $f : X \to \mathbb{R}$ с нормой $\|f\| = \sup_{\|x\| \leq 1} |f(x)|$.

Утверждение. На конечномерном нормированном пространстве любой линейный функционал непрерывен. На любом бесконечномерном нормированном пространстве существует неограниченный линейный функционал.

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 1. 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.

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 2. 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 1. (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 2. (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*

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

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 3. 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 $U \subset X$ — 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.

Примеры вычисления производных. 1) 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.

2) Пусть H — евклидово пространство, $\langle \cdot, \cdot \rangle$ — скалярное произведение, $f(x) = \langle x, x \rangle$. Тогда $f(x + h) = f(x) + 2\langle x, h \rangle + \langle h, h \rangle$. Отображение $h \mapsto 2\langle x, h \rangle$ линейно и непрерывно (в силу неравенства Коши – Буняковского), $\langle h, h \rangle = \|h\|^2 = o(\|h\|)$, $\|h\| \to 0$. Наконец, отображение $x \mapsto 2\langle x, \cdot \rangle$ непрерывно. В самом деле,

$$\|2\langle x, \cdot \rangle - 2\langle x_0, \cdot \rangle\| = \sup_{\|h\| \le 1} |2\langle x, h \rangle - 2\langle x_0, h \rangle| \le 2\|x - x_0\|.$$

Напоминание (теорема Лагранжа). Пусть $f : [a, b] \to \mathbb{R}$ непрерывна на $[a, b]$ и дифференцируема на (a, b). Тогда существует точка $\xi \in (a, b)$ такая, что $f(b) - f(a) = f'(\xi)(b - a)$.

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 3. *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*

$$\|F(x_1) - F(x_0)\| \le \sup_{x \in [x_0,\, x_1]} \|F'(x)[x_1 - x_0]\|. \qquad (1)$$

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,

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

$\square$

Задача 4. 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 4. *Если отображение непрерывно дифференцируемо в т. x_0, то оно дифференцируемо по Фреше в т. x_0.*

Доказательство. Нам нужно показать, что

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

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

$$\|F(x_0 + h) - F(x_0) - F'(x_0)[h]\| = \|G(x_0 + h) - G(x_0)\| \overset{(1)}{\le} \sup_{x \in [x_0,\, x_0 + h]} \|G'(x)[h]\| \le$$

$$\leq \sup_{x\in[x_0,\, x_0+h]} \|F'(x) - F'(x_0)\| \cdot \|h\|.$$

Так как F' непрерывно в x_0 и $x \in [x_0,\, x_0+h]$, то $\sup_{x\in[x_0,\, x_0+h]} \|F'(x) - F'(x_0)\| \cdot \|h\| = o(\|h\|)$, $\|h\| \to 0$. $\square$

Пример. Пусть

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Тогда f дифференцируемо по Фреше в каждой точке $x \in \mathbb{R}$, но не является непрерывно дифференцируемым в точке $x = 0$.

2 Оператор Немыцкого

Напомним, что $C([t_0,\, t_1], \mathbb{R}^n)$ — это пространство непрерывных функций $x : [t_0,\, t_1] \to \mathbb{R}^n$; на этом пространстве вводится норма $\|x\|_C = \max_{t\in[t_0,\, t_1]} |x(t)|$ (здесь $|x(t)|$ — длина вектора $x(t)$ в пространстве $\mathbb{R}^n$).

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

Теорема 1. *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*

$$N'_\varphi(x)[h](t) = \varphi'_\xi(t,\, x(t))h(t); \tag{2}$$

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 (2). 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

$$\|\varphi'_{j,\xi}(t,\, x(t)+w)-\varphi'_{j,\xi}(t,\, x(t))\|<\varepsilon. \tag{3}$$

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 (3). $\square$

Задача 5. Пусть X — пространство непрерывных функций на $[0,\, 1]$ с нормой $\|x\| = \int\limits_0^1 |x(t)|\, dt$, $T : X \to X$, $Tx(t) = x^2(t)$. Показать, что T дифференцируем по Гато, но не дифференцируем по Фреше.

Пространство $C^1([t_0,\, t_1],\, \mathbb{R}^n)$ определяется как пространство непрерывно дифференцируемых функций $x : [t_0,\, t_1] \to \mathbb{R}^n$, снабженное нормой $\|x\|_{C^1} = \max\{\|x\|_C,\, \|\dot{x}\|_C\}$.

Пусть $L : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ непрерывно. Обозначим элементы $[t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^n$ через $(t,\, \xi_1,\, \ldots,\, \xi_n,\, \eta_1,\, \ldots,\, \eta_n)$. Предположим, что частные производные $L_{\xi_i} : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ и $L_{\eta_i} : [t_0,\, t_1] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ непрерывны для любого $i = 1,\, \ldots,\, n$. Рассмотрим отображение $\mathcal{L} : C^1([t_0,\, t_1],\, \mathbb{R}^n) \to \mathbb{R}$,

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

Тогда $\mathcal{L}$ дифференцируемо по Фреше. В самом деле, $\mathcal{L} = f \circ N_L \circ A$, где

$$A : C^1([t_0,\, t_1],\, \mathbb{R}^n) \to C([t_0,\, t_1],\, \mathbb{R}^{2n}), \quad A(x_1,\, \ldots,\, x_n) = (x_1,\, \ldots,\, x_n,\, \dot{x}_1,\, \ldots \dot{x}_n),$$

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

Отображения A и f линейны и непрерывны, оператор Немыцкого N_L непрерывно дифференцируем. Значит, $\mathcal{L}$ дифференцируемо по Фреше.

При этом

$$\mathcal{L}'(x)[h] = \int_{t_0}^{t_1} (L_\xi(t,\, x(t),\, \dot{x}(t))h(t) + L_\eta(t,\, x(t),\, \dot{x}(t))\dot{h}(t))\, dt. \tag{4}$$

Задача 6. Показать, что отображение $\mathcal{L}$ непрерывно дифференцируемо.

3 Гладкая задача на минимум без ограничений.

Пусть X — нормированное пространство, $U \subset X$ — открытое множество, $f : U \to \mathbb{R}$. Рассмотрим задачу поиска точки минимума функции f на U:

$$f(x) \to \min, \quad x \in U. \tag{5}$$

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 5. *Пусть $\hat{x}$ — точка локального минимума в задаче (5), и пусть f имеет вариацию по Лагранжу в точке $\hat{x}$. Тогда $f'(\hat{x}) = 0$.*

Доказательство. Если $f'(\hat{x}) \neq 0$, то существует вектор $h \in X$ такой, что $f'(\hat{x})[h] \neq 0$. Пусть $\varphi(t) = f(\hat{x} + th)$. Тогда $t = 0$ является точкой локального минимума функции φ и поэтому $\varphi'(0) = 0$. Но $\varphi'(0) = f'(\hat{x})[h]$ — противоречие. □

Хорошо известно, что равенство нулю вариации по Лагранжу не является достаточным условием для локального минимума (вспомним функцию $f(x) = x^3$). Позже мы получим необходимые и достаточные условия локального минимума в терминах вторых производных. А сейчас будет получено достаточное условие глобального минимума.

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 $f((1-\lambda)x + \lambda y) \le (1-\lambda)f(x) + \lambda f(y)$.

Предложение 6. *Пусть X — линейное нормированное пространство, $U \subset X$ — выпуклое открытое подмножество, $f : U \to \mathbb{R}$ — выпуклая функция, $\hat{x} \in U$, $f'(\hat{x}) = 0$. Тогда $\hat{x}$ — точка глобального минимума.*

Доказательство. Пусть $\hat{x}$ не является точкой минимума. Тогда существует вектор $h \in X$ 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 $f'(\hat{x})[h] \le f(\hat{x}+h) - f(\hat{x}) < 0$. $\square$

4 Простейшая задача вариационного исчисления

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).

Пусть 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

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

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. Тогда, как говорилось выше, отображение $\mathcal{L}$ непрерывно дифференцируемо на $C^1([t_0,\, t_1],\, \mathbb{R}^n)$.

Далее будем обозначать: $L_{\xi_i} =: L_{x_i}$, $L_{\eta_i} =: L_{\dot{x}_i}$.

Рассмотрим следующую задачу на минимум:

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt \to \inf, \quad x(t_0) = x_{(0)}, \quad x(t_1) = x_{(1)}. \tag{6}$$

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 (6), если она допустимая и существует $\varepsilon > 0$ такое, что для любой допустимой функции x такой, что $\|x - \hat{x}\|_{C^1} < \varepsilon$ выполнено $\mathcal{L}(x) \geq \mathcal{L}(\hat{x})$.

Определим пространство $C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n)$, состоящее из функций $g \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$ таких, что $g(t_0) = g(t_1) = 0$.

Пусть $\hat{x}$ — допустимая функция. Рассмотрим функционал $\tilde{\mathcal{L}} : C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n) \to \mathbb{R}$, определенный равенством

$$\tilde{\mathcal{L}}(g) = \mathcal{L}(\hat{x} + g). \tag{7}$$

Тогда

$$\tilde{\mathcal{L}}(g) = \int\limits_{t_0}^{t_1} \tilde{L}(t,\, g(t),\, \dot{g}(t))\, dt, \quad \tilde{L}(t,\, \xi,\, \eta) = L(t,\, \hat{x}(t) + \xi,\, \dot{\hat{x}}(t) + \eta);$$

$\hat{x}$ является точкой локального минимума в (6) тогда и только тогда, когда 0 является точкой локального минимума в задаче

$$\tilde{\mathcal{L}}(g) \to \inf, \quad g \in C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n). \tag{8}$$

Это задача без ограничений на линейном нормированном пространстве.

Предложение 7. *Пусть $\hat{x}$ — точка локального минимума в задаче (6). Тогда для любого $h \in C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n)$ выполнено*

$$\int\limits_{t_0}^{t_1} \Big(L_x(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))h(t) + L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))\dot{h}(t) \Big)\, dt = 0, \tag{9}$$

т.е.

$$\int\limits_{t_0}^{t_1} \left(\sum_{j=1}^{n} L_{x_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))h_j(t) + \sum_{j=1}^{n} L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))\dot{h}_j(t) \right) dt = 0. \tag{10}$$

Доказательство. Имеем $\tilde{\mathcal{L}}'(0)[h] = 0$ для любого $h \in C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n)$. Так как $\tilde{L}_{\xi_i}(t,\, \xi,\, \eta) = L_{x_i}(t,\, \hat{x}(t) + \xi,\, \dot{\hat{x}}(t) + \eta)$, $\tilde{L}_{\eta_i}(t,\, \xi,\, \eta) = L_{\dot{x}_i}(t,\, \hat{x}(t) + \xi,\, \dot{\hat{x}}(t) + \eta)$ и эти функции непрерывны по совокупности аргументов $(t,\, \xi,\, \eta)$, то функционал $\tilde{\mathcal{L}}$ непрерывно дифференцируем в 0, и в силу (4)

$$\tilde{\mathcal{L}}'(0)[h] = \int\limits_{t_0}^{t_1} \Big(L_x(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))h(t) + L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))\dot{h}(t) \Big)\, dt.$$

Отсюда следует (10). □

Оказывается, (10) эквивалентно некоторой системе обыкновенных дифференциальных уравнений, называющейся уравнениями Эйлера – Лагранжа.

Сначала напомним одно из эквивалентных определений абсолютно непрерывной функции.

Определение. *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:

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

Предложение 8. *Равенство (10) для любого $h \in C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n)$ выполнено выполнено тогда и только тогда, когда функция $t \mapsto L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))$ абсолютно непрерывна и*

$$-\frac{d}{dt}L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) + L_{x_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) = 0, \quad 1 \le j \le n; \tag{11}$$

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. Пусть выполнено (10) для любого $h \in C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n)$. Положим $B_j(t) = \int\limits_{t_0}^{t} L_{x_j}(s,\, \hat{x}(s),\, \dot{\hat{x}}(s))\, ds$. Проинтегрируем по частям в (10) $\sum\limits_{j=1}^{n} L_{x_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))h_j(t)$, учитывая, что $h_j(t_0) = h_j(t_1) = 0$, и получим:

$$\int\limits_{t_0}^{t_1} \sum_{j=1}^{n}(L_{\dot{x}_j}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t)) - B_j(t))\dot{h}_j(t)\, dt = 0. \tag{12}$$

Нам понадобится следующий вспомогательный факт.

Лемма Дюбуа-Реймона. Пусть $B : [t_0,\, t_1] \to \mathbb{R}$ — непрерывная функция такая, что для любой функции $z \in C^1[t_0,\, t_1]$ такой, что $z(t_0) = z(t_1) = 0$, выполнено $\int\limits_{t_0}^{t_1} B(t)\dot{z}(t)\, dt = 0$. Тогда $B \equiv \text{const}$.

Доказательство. Пусть функция B не является константой. Тогда существуют точки $\tau_1,\, \tau_2 \in (t_0,\, t_1)$ такие, что $B(\tau_1) < B(\tau_2)$. Так как функция B непрерывна, то существуют $c \in \mathbb{R}$, $\varepsilon > 0$ и $\delta > 0$ такие, что $B(t) \le c - \varepsilon$ для любого $t \in O_\delta(\tau_1)$ и $B(t) \ge c + \varepsilon$ для любого $t \in O_\delta(\tau_2)$. Пусть

$$g(t) = \begin{cases} \delta - |t - \tau_1|, & |t - \tau_1| < \delta, \\ |t - \tau_2| - \delta, & |t - \tau_2| < \delta, \\ 0, \quad \text{иначе}. \end{cases}$$

Тогда $\int\limits_{t_0}^{t_1} g(s)\,ds = 0$. Положим $z(t) = \int\limits_{t_0}^{t} g(s)\,ds$. Тогда $z \in C^1[t_0,\, t_1]$, $z(t_0) = z(t_1) = 0$. С другой стороны,

$$\int\limits_{t_0}^{t_1} B(t)\dot z(t)\,dt = \int\limits_{O_\delta(\tau_1)} B(t)g(t)\,dt + \int\limits_{O_\delta(\tau_2)} B(t)g(t)\,dt \le$$

$$\le (c-\varepsilon)\int\limits_{O_\delta(\tau_1)} (\delta-|t-\tau_1|)\,dt+(c+\varepsilon)\int\limits_{O_\delta(\tau_2)} (-\delta+|t-\tau_2|)\,dt = -2\varepsilon\int\limits_{O_\delta(0)} (\delta-|t|)\,dt < 0.$$

Лемма Дюбуа-Реймона доказана. $\diamond$

Из (12) и леммы Дюбуа-Реймона следует, что

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

Значит, $L_{\dot x_j}(t,\, \hat x(t),\, \dot{\hat x}(t))$ абсолютно непрерывны и выполнено (11).

Обратно, пусть $L_{\dot x_j}(t,\, \hat x(t),\, \dot{\hat x}(t))$ абсолютно непрерывны и выполнено (11). Тогда

$$\int\limits_{t_0}^{t_1} \sum_{j=1}^n \left(-\frac{d}{dt}L_{\dot x_j}(t,\, \hat x(t),\, \dot{\hat x}(t)) + L_{x_j}(t,\, \hat x(t),\, \dot{\hat x}(t))\right) h_j(t)\,dt = 0, \quad h \in C^1_{0,0}([t_0,\, t_1], \mathbb{R}^n).$$

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 (10). $\square$

Итак, доказана

Теорема 2. *Пусть $\hat x$ — точка локального минимума в (6). Тогда выполнены уравнения Эйлера – Лагранжа (11).*

Допустимая функция $\hat x$, удовлетворяющая (11), называется допустимой экстремалью.

Ближе к концу курса мы изучим, когда допустимая экстремаль будет точкой локального минимума (будут получены необходимые и достаточные условия второго порядка). Сейчас мы получим достаточное условие глобального минимума.

Предложение 9. *Пусть функционал $\mathcal{L}$ выпуклый, $\hat{x}$ — допустимая экстремаль. Тогда $\hat{x}$ является точкой глобального минимума в задаче (6).*

Доказательство. Если $\hat{x}$ — допустимая экстремаль, то $\tilde{\mathcal{L}}'(0) = 0$. Так как $\mathcal{L}$ выпуклый, то $\tilde{\mathcal{L}}$ выпуклый (см. (7)). Значит, 0 — точка глобального минимума $\tilde{\mathcal{L}}$ на $C^1_{0,0}([t_0,\, t_1],\, \mathbb{R}^n)$. Отсюда получаем, что $\hat{x}$ — точка глобального минимума в (6). □

Предложение 10. *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.*

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

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

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

□

Гармонический осциллятор. Рассмотрим задачу

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

Тогда $L_{\dot{x}} = 2\dot{x}$, $L_x = -2x$; уравнение Эйлера имеет вид $-\frac{d}{dt}(2\dot{x}) - 2x = 0$, т.е. $\ddot{x} + x = 0$.

Заметим, что $\hat{x} = 0$ является допустимой экстремалью. Выясним, является ли она точкой локального или глобального минимума. Для этого используем следующий прием (когда будем изучать условия второго порядка в простейшей задаче вариационного исчисления, будет описано его обобщение).

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,

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

Подберем ω так, чтобы $\dot{x}^2 - x^2 - \dot{\omega}x^2 - 2\omega x\dot{x}$ было полным квадратом: $\dot{x}^2 - x^2 - \dot{\omega}x^2 - 2\omega x\dot{x} = (\dot{x} - \omega x)^2$, т.е. $-1 - \dot{\omega} = \omega^2$. (Тогда $\int\limits_0^{T_0} (\dot{x}^2 - x^2)\,dt = \int\limits_0^{T_0} (\dot{x} - \omega x)^2\,dt \geq 0$.) Получаем, что $\omega = \operatorname{ctg}(t - t_*)$.

Задача 7. 1) Пусть $T_0 > \pi$, $x(t) = c\sin\frac{\pi t}{T_0}$. Показать, что $\mathcal{L}(x) < 0$. Почему проведенные выше рассуждения не проходят при $T_0 > \pi$ и проходят при $T_0 < \pi$? 2) Показать, что $\int\limits_0^{\pi} (\dot{x}^2 - x^2)\,dt = \int\limits_0^{\pi} (\dot{x} - x\cdot\operatorname{ctg} t)^2\,dt \geq 0$.

Вывод. Если $T_0 \leq \pi$, то 0 — точка глобального минимума. Если $T_0 > \pi$, то 0 не является точкой локального минимума.

Задача 8. 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

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

Задача 9. 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.

5 Задача Больца

Пусть $L : [t_0,\,t_1]\times\mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ — такая же, как в предыдущем параграфе; $l : \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ 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 $\frac{\partial l}{\partial x_j(t_1)}$.

Функционал $\mathcal{L}$ непрерывно дифференцируем и

$$\mathcal{L}'(x)[h] = \int\limits_{t_0}^{t_1} \left(\sum_{j=1}^{n} L_{\dot{x}_j}(t,\, x(t),\, \dot{x}(t))\dot{h}_j(t) + \sum_{j=1}^{n} L_{x_j}(t,\, x(t),\, \dot{x}(t))h_j(t) \right) dt+$$

$$+ \sum_{j=1}^{n} \frac{\partial l}{\partial x_j(t_0)} h_j(t_0) + \sum_{j=1}^{n} \frac{\partial l}{\partial x_j(t_1)} h_j(t_1) =$$

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

Рассмотрим задачу Больца:

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt + l(x(t_0),\, x(t_1)) \to \inf. \tag{13}$$

Функция $\hat{x}$ называется точкой локального минимума в задаче (13), 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 $\mathcal{L}(x) \geq \mathcal{L}(\hat{x})$.

Если $\hat{x}$ — точка локального минимума, то $\mathcal{L}'(\hat{x})[h] = 0$ для любого $h \in C^1([t_0,\, t_1],\, \mathbb{R}^n)$. Значит,

$$\begin{gathered} \int\limits_{t_0}^{t_1} \left(\sum_{j=1}^{n} L_{\dot{x}_j}(t,\, x(t),\, \dot{x}(t))\dot{h}_j(t) + \sum_{j=1}^{n} L_{x_j}(t,\, x(t),\, \dot{x}(t))h_j(t) \right) dt+ \\ + \sum_{j=1}^{n} \frac{\partial l}{\partial x_j(t_0)} h_j(t_0) + \sum_{j=1}^{n} \frac{\partial l}{\partial x_j(t_1)} h_j(t_1) = 0 \end{gathered} \tag{14}$$

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

Предложение 11. *Условие (14) эквивалентно выполнению уравнений Эйлера – Лагранжа (11) 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 (14). Подставив произвольную функцию $h \in C^1_{0,0}([t_0,\, t_1],\, \mathbb{R}^n)$, получим, что

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

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

Теперь докажем условия трансверсальности. Интегрируя по частям левую часть (14), получаем, что

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

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 $i \ne j$, $h_j(t_0) = 1$, $h_j(t_1) = 0$. 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 $i \ne j$, $h_j(t_0) = 0$, $h_j(t_1) = 1$, то получим второе условие трансверсальности.

Обратно, пусть выполнены уравнения Эйлера и условие трансверсальности. Тогда

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

$$+ \sum_{j=1}^{n} \left(-L_{\dot{x}_j}(t_0,\, \hat{x}(t_0),\, \dot{\hat{x}}(t_0)) + \frac{\partial l}{\partial x_j(t_0)}(\hat{x}(t_0),\, \hat{x}(t_1)) \right) h_j(t_0) +$$

$$+ \sum_{j=1}^{n} \left(L_{\dot{x}_j}(t_1,\, \hat{x}(t_1),\, \dot{\hat{x}}(t_1)) + \frac{\partial l}{\partial x_j(t_1)}(\hat{x}(t_0),\, \hat{x}(t_1)) \right) h_j(t_1) = 0.$$

Проинтегрировав по частям, получаем (14). □

Отсюда получаем следующий результат.

Теорема 3. *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.

Если функционал $\mathcal{L}$ выпуклый, то допустимая экстремаль является точкой глобального минимума.

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

Предложение 12. *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.

6 Гладкая задача с ограничениями типа равенств

6.1 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

$$\begin{cases} f_0(x) \to \inf, \\ F(x) = 0. \end{cases} \tag{15}$$

Точка $x \in U$ называется допустимой для задачи (15), 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 $f_0(x) \geq f_0(\hat{x})$.

Если $Y = \mathbb{R}^m$, $F(x) = (f_1(x), \ldots, f_m(x))$, то (15) переписывается в виде

$$\begin{cases} f_0(x) \to \inf, \\ f_j(x) = 0, \quad 1 \leq j \leq m. \end{cases} \tag{16}$$

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

Предложение 13. *Пусть $X = \mathbb{R}^n$, $\hat{x}$ — точка локального минимума в (16), $f_j \in C^1(U)$, $0 \leq j \leq m$, $m \leq n$, матрица $\left(\frac{\partial f_j}{\partial x_i}(\hat{x})\right)_{1 \leq i \leq n,\, 1 \leq j \leq m}$ имеет максимальный ранг. Тогда существуют $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ 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 (15) записывается в виде

$$\mathcal{L}(x) = \lambda_0 f_0(x) + y^*(F(x)), \tag{17}$$

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 $y^*(y_1, \ldots, y_m) = \sum\limits_{j=1}^{m} \lambda_j y_j$, 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 (15), f_0, F непрерывно 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 (17). При этом, если оператор $F'(\hat{x})$ сюръективный, то можно взять $\lambda_0 \neq 0$.*

Приведем два примера.

Задача 10. Пусть $A : l_2 \to l_2$,

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

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

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

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 11. 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).

6.2 Касательные векторы

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

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

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

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

(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.)

Задача 12. Найти

1. $T_{(0,0)}M$, где $M = \{(x,\, y) \in \mathbb{R}^2 :\ y^3 = x^2\}$,
2. $T_{(0,0)}M$, где $M = \{(x,\, y) \in \mathbb{R}^2 :\ y^2 = \sin^2 x\}$,
3. $T_{(0,0,0)}M$, где $M = \{(x,\, y,\, z) \in \mathbb{R}^3 :\ (z - x^2 - y^2)(x^2 - y^3) = 0\}$.

Задача 13. Пусть $M = \cup_{n \in \mathbb{N}} \{(x,\, y) \in \mathbb{R}^2 :\ \ y = x/n,\ -1/n \leq x \leq 1/n\}$. Найти $T_{(0,0)}M$.

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

$$\begin{cases} f_0(x) \to \inf, \\ x \in M. \end{cases} \tag{18}$$

Точка $\hat{x} \in M$ называется точкой локального минимума в задаче (18), 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 14. *Пусть $\hat{x} \in M$ — точка локального минимума в задаче (18), 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 14. 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.

6.3 Фактор-пространства

Пусть X — банахово пространство, $Y \subset X$ — замкнутое подпространство. Напомним, что фактор-пространство X/Y — это множество $\{x + Y :$

$x \in X\}$. Заметим, что $x + Y = z + Y$ тогда и только тогда, когда $x - z \in Y$. Операции сложения и умножения на число задаются равенствами $(x+Y)+(z+Y) := x+z+Y$, $\alpha(x+Y) := \alpha x+Y$. Они задают структуру линейного пространства на X/Y.

Теперь введем норму на X/Y равенством

$$\|x + Y\|_{X/Y} = \mathrm{dist}\,(x,\, Y).$$

Сначала отметим, что эта величина корректно определена, т.е. если $x + Y = \tilde{x}+Y$, то $\|x+Y\|_{X/Y} = \|\tilde{x}+Y\|_{X/Y}$. В самом деле, для любого $y \in Y$

$$\|\tilde{x} - y\| = \|x - (x - \tilde{x} + y)\| \geq \mathrm{dist}\,(x,\, Y),$$

так что $\mathrm{dist}\,(\tilde{x},\, Y) \geq \mathrm{dist}\,(x,\, Y)$. Меняем местами x и $\tilde{x}$ и получаем $\mathrm{dist}\,(x,\, Y) = \mathrm{dist}\,(\tilde{x},\, Y)$.

Теперь проверим, что $\|\cdot\|_{X/Y}$ — норма. Ясно, что это неотрицательная величина. Если $x + Y \neq 0$, то $x \notin Y$; так как Y замкнуто, то $\mathrm{dist}\,(x,\, Y) > 0$. Равенство $\|\alpha x + Y\| = |\alpha| \cdot \|x + Y\|$ легко следует из определения. Докажем неравенство треугольника. Пусть $x_1,\, x_2 \in X$. Для произвольных $y_1,\, y_2 \in Y$ выполнено

$$\mathrm{dist}\,(x_1 + x_2,\, Y) \leq \|x_1 + x_2 - y_1 - y_2\| \leq \|x_1 - y_1\| + \|x_2 - y_2\|.$$

Взяв inf правой части по $y_1,\, y_2$, получаем

$$\mathrm{dist}\,(x_1 + x_2,\, Y) \leq \mathrm{dist}\,(x_1,\, Y) + \mathrm{dist}\,(x_2,\, Y).$$

Теорема 5. *Пространство X/Y полно.*

Доказательство. Пусть последовательность $\{x_n+Y\}_{n\in\mathbb{N}}$ фундаментальна в X/Y. Тогда существует подпоследовательность $\{n_k\}_{k\in\mathbb{N}}$ такая, что $\|x_{n_{k+1}} - x_{n_k} + Y\|_{X/Y} < 2^{-k}$, $k \in \mathbb{N}$. Значит, для каждого $k \in \mathbb{N}$ существует вектор $z_k \in Y$ такой, что $\|x_{n_{k+1}} - x_{n_k} - z_k\|_X < 2^{-k}$. Положим

$$\tilde{x}_{n_1} = x_{n_1}, \quad \tilde{x}_{n_k} = x_{n_k} - \sum_{j=1}^{k-1} z_j. \tag{19}$$

Тогда $\tilde{x}_{n_{k+1}} - \tilde{x}_{n_k} = x_{n_{k+1}} - x_{n_k} - z_k$, откуда $\|\tilde{x}_{n_{k+1}} - \tilde{x}_{n_k}\|_X \leq 2^{-k}$. Поэтому $\{\tilde{x}_{n_k}\}_{k\in\mathbb{N}}$ фундаментальна. Пространство X банахово; значит, существует вектор $\hat{x} \in X$ такой, что $\tilde{x}_{n_k} \underset{k\to\infty}{\to} \hat{x}$ в X. Отсюда следует, что $\tilde{x}_{n_k} + Y \underset{k\to\infty}{\to} \hat{x} + Y$ в X/Y. Из (19) следует, что $x_{n_k} + Y = \tilde{x}_{n_k} + Y$. Значит, $x_{n_k} + Y \underset{k\to\infty}{\to} \hat{x} + Y$ в X/Y. □

6.4 Лемма о правом обратном

Будем далее обозначать через B_X замкнутый единичный шар пространства X.

Сначала напомним теорему Банаха об обратном операторе.

Теорема 6. *Пусть X, Y — банаховы пространства, $A : X \to Y$ — линейный непрерывный биективный оператор. Тогда оператор $A^{-1} : Y \to X$ непрерывен.*

Отсюда получаем следующее утверждение.

Теорема 7. *Пусть X, Y — банаховы пространства, $A : X \to Y$ — линейный непрерывный сюръективный оператор. Тогда существует $\delta > 0$ такое, что $A(B_X) \supset \delta \cdot B_Y$.*

Доказательство. Пусть $Z = \ker A$. Определим оператор $T : X/Z \to Y$ по формуле $T(x + Z) = A(x)$ (это отображение корректно определено, линейно и биективно).

Проверим, что

$$T(B_{X/Z}) \subset A(B_X); \tag{20}$$

отсюда и из ограниченности оператора A будет следовать, что T ограничен.

Пусть $\|x + Z\|_{X/Z} < 1$. Тогда существует вектор $z \in Z$ такой, что $\|x - z\|_X < 1$. Значит, $T(x + Z) = A(x - z) \subset A(B_X)$.

По теореме Банаха об обратном операторе, T^{-1} ограничен. Значит, существует $\delta > 0$ такое, что $T^{-1}(\delta B_Y) \subset B_{X/Z}$; таким образом, $\delta B_Y \subset T(B_{X/Z}) \overset{(20)}{\subset} A(B_X)$. □

Теперь докажем лемму о правом обратном.

Лемма. *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 7, 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 $Ry = \frac{\|y\|}{\delta} x\left(\frac{\delta y}{\|y\|}\right)$, $y \ne 0$; $R(0) = 0$. □

6.5 Теорема Люстерника и теорема о касательном пространстве

Сначала отметим следующее свойство непрерывно дифференцируемых отображений.

Предложение 15. *Пусть X, Y — нормированные пространства, $U \subset X$ — открытое множество, $F : U \to Y$, $x_0 \in U$, F непрерывно дифференцируемо в точке x_0. Тогда для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любых x_1, $x_2 \in U$ таких, что $\|x_1 - x_0\| < \delta$, $\|x_2 - x_0\| < \delta$, выполнено*

$$\|F(x_1) - F(x_2) - F'(x_0)(x_1 - x_2)\| \le \varepsilon \|x_1 - x_2\|.$$

Доказательство. Пусть $G(x) = F(x) - F'(x_0)[x]$. Тогда G дифференцируемо по Гато в окрестности точки x_0 и $G'(x) = F'(x) - F'(x_0)$. Так как отображение F' непрерывно в точке x_0, то для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любого $x \in O_\delta(x_0)$ выполнено $\|F'(x) - F'(x_0)\|_{L(X,Y)} \le \varepsilon$. Значит, если $\|x_1 - x_0\| < \delta$, $\|x_2 - x_0\| < \delta$, то

$$\|F(x_1) - F(x_2) - F'(x_0)(x_1 - x_2)\| = \|G(x_1) - G(x_2)\| \overset{(1)}{\le}$$

$$\le \sup_{x \in [x_1, x_2]} \|G'(x)\| \cdot \|x_1 - x_2\| = \sup_{x \in [x_1, x_2]} \|F'(x) - F'(x_0)\| \cdot \|x_1 - x_2\| \le \varepsilon \|x_1 - x_2\|.$$

□

Следствие 1. *Если F непрерывно дифференцируемо в точке x_0, то F непрерывно в некоторой окрестности точки x_0.*

Теперь докажем теорему Люстерника.

Теорема 8. *Пусть X, Y — банаховы пространства, $U \subset X$ — открытое множество, $x_0 \in U$, $F : U \to Y$ непрерывно 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

$$\|x - \varphi(x)\| \le K \|F(x)\|. \tag{21}$$

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

$$\|Ry\| \leq C\|y\|. \tag{22}$$

Итак, записываем итерационный процесс:

$$x_1 = x, \quad x_{n+1} = x_n - RF(x_n). \tag{23}$$

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 (23) оператор $F'(x_0)$:

$$F'(x_0)(x_{n+1} - x_n) + F(x_n) = 0. \tag{24}$$

Возьмем произвольное $\varepsilon > 0$ и выберем для него $\delta > 0$ в соответствии с предложением 15. 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 (24), получаем

$$\|F(x_{n+1})\| \leq \varepsilon\|x_{n+1} - x_n\|. \tag{25}$$

Снова из (23) получаем:

$$\|x_{n+1}-x_n\| = \|RF(x_n)\| \le C\|F(x_n)\| \overset{(25)}{\le} C\varepsilon\|x_n - x_{n-1}\| \tag{26}$$

(последнее неравенство выполнено при $n \ge 2$).

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

$$\|x_{n+1}-x_n\| \le \frac{1}{2}\|x_n - x_{n-1}\|. \tag{27}$$

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

$$\begin{aligned} \|x_{n+k}-x_n\| \le \|x_{n+k}-x_{n+k-1}\| + \cdots + \|x_{n+1}-x_n\| \overset{(27)}{\le} \\ \le (2^{-n-k+2}+\cdots+2^{-n+1})\|x_2-x_1\| \le 2^{-n+2}\|x_2-x_1\|. \end{aligned} \tag{28}$$

Величина $\|x_2-x_1\|$ оценивается с помощью (26): $\|x_2-x_1\| \le C\|F(x_1)\|$; так как отображение F непрерывно в точке x_0 (см. следствие 1), то правая часть будет мала, если $x_1 = x$ будет принадлежать малой окрестности точки x_0. А значит, учитывая (28), 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 (28) 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 (28) и получаем:

$$\|x_n - x_1\| \le 4\|x_2-x_1\| \overset{(26)}{\le} 4C\|F(x_1)\|;$$

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

Из (25) и (28) следует, что $F(x_n) \underset{n\to\infty}{\to} 0$; с другой стороны, из следствия 1 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. $\varphi(x) \in M$. □

Теперь докажем теорему о касательном пространстве.

Теорема 9. *Пусть X, Y — банаховы пространства, $U \subset X$ — открытое множество, $x_0 \in U$, $F : U \to Y$ непрерывно 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

$$x_0 + th + r(t) \in M, \tag{29}$$

$$\|r(t)\| \leq K\|F(x_0 + th)\|. \tag{30}$$

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, $\|r(t)\| \overset{(30)}{=} o(t)$, $t \to 0$. В силу (29), 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.

6.6 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 10. *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 $x^* \in (\ker A)^\perp$. Определим функционал $f : X/\ker A \to \mathbb{R}$ равенством $f(x + \ker A) = x^*(x)$. Так как $x^* \in (\ker A)^\perp$, то f корректно определено. Отображение f линейно. Так как для любого $x \in X$ существует вектор $z \in \ker A$ такой, что $\|x - z\| \le 2\|x + \ker A\|$, то

$$|f(x + \ker A)| = |x^*(x - z)| \le \|x^*\| \cdot \|x - z\| \le 2\|x^*\| \cdot \|x + \ker A\|,$$

откуда $\|f\| \le 2\|x^*\|$. Значит, $f \in (X/\ker A)^*$.

Мы уже определяли оператор $T : X/\ker A \to Y$, $T(x + \ker A) = A(x)$ и показывали, что это линейная непрерывная биекция.

Имеем: $x^*(x) = f(x + \ker A) = f \circ T^{-1}(Ax)$; остается заметить, что $f \circ T^{-1} \in Y^*$. $\square$

6.7 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, $f_0'(\hat{x}) \in \operatorname{Im}(F'(\hat{x}))^*$, 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 $\operatorname{Im} F'(x_0) \ne Y$ — замкнутое подпространство. Нам понадобится лемма о нетривиальном аннуляторе.

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

Доказательство. Пусть $\xi \in X$, $\operatorname{dist}(\xi, L) = 1$. Положим $z^*(x + \alpha\xi) = \alpha$, если $x \in L$, $\alpha \in \mathbb{R}$. Тогда z^* — линейный функционал на $\operatorname{span}(L \cup \{\xi\})$, $z^*|_L = 0$. Докажем, что z^* непрерывен. В самом деле,

$$\|x + \alpha\xi\| \ge \operatorname{dist}(\alpha\xi, L) = |\alpha| \operatorname{dist}(\xi, L) = |\alpha|,$$

то есть $\|z^*\| \le 1$.

По теореме Хана–Банаха, существует продолжение z^* на X с сохранением нормы. Это и будет искомый функционал. $\square$

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

7 Изопериметрическая задача

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

$$\begin{cases} \int\limits_{t_0}^{t_1} L_0(t, x(t), \dot{x}(t))\, dt \to \inf, \\ \int\limits_{t_0}^{t_1} L_j(t, x(t), \dot{x}(t))\, dt = c_j, \quad 1 \le j \le m, \\ x(t_0) = x_{(0)},\ x(t_1) = x_{(1)}, \\ x \in C^1([t_0, t_1], \mathbb{R}^n). \end{cases} \tag{31}$$

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 (31), 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 (31), 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 (31) 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 (31). 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 11. *Пусть $\hat{x}$ — точка локального минимума в задаче (31). 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 (31) дополнительно написать ограничения $x(t) \in U$, $\dot{x}(t) \in V$, то необходимые условия локального минимума формулируются так же.

Задача 15. Пусть $l > 0$. Доказать, что допустимые экстремали в задаче

$$\int\limits_0^1 (y\dot{x} - x\dot{y})\, dt \to \max, \quad \int\limits_0^1 \sqrt{\dot{x}^2 + \dot{y}^2}\, dt = l, \quad x(0) = x(1) = y(0) = y(1) = 0, \quad \dot{x}^2 + \dot{y}^2 > 0 \qquad (32)$$

являются параметризацией окружности.

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 (32) будет сведена к эквивалентной задаче, у которой существование минимума уже будет нетрудно доказать.

8 Гладкая задача с ограничениями типа равенств и неравенств

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

$$\begin{cases} f_0(x) \to \inf, \\ f_j(x) \le 0, \quad 1 \le j \le m, \\ F(x) = 0, \\ x \in U. \end{cases} \tag{33}$$

Точка $x \in U$ называется допустимой для задачи (33), если $f_j(x) \le 0$ для любого $j = 1, \dots, m$ и $F(x) = 0$. Допустимая точка $\hat{x}$ называется точкой локального минимума в задаче (33), если существует $\varepsilon > 0$ такое, что для любой допустимой точки x такой, что $\|x - \hat{x}\| < \varepsilon$, выполнено $f_0(x) \ge f_0(\hat{x})$.

Теорема 12. *Пусть пространства X, Y полны, отображения $f_0, \dots, f_m$, F непрерывно дифференцируемы в точке $\hat{x}$. Пусть $\operatorname{Im} F'(\hat{x})$ замкнут. Тогда существуют числа $\lambda_0, \dots, \lambda_m \in \mathbb{R}$ и функционал $y^* \in Y^*$, одновременно не равные нулю, удовлетворяющие следующим условиям:*

1. *$\lambda_j \ge 0$, $0 \le j \le m$ (условие неотрицательности);*
2. *$\lambda_j f_j(\hat{x}) = 0$, $1 \le j \le m$ (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

$$\sum_{j=0}^{m} \lambda_j f_j'(\hat{x})[h] + y^*(F'(\hat{x})[h]) = 0. \tag{34}$$

Перед тем, как доказать теорему 12, сформулируем частный случай теоремы отделимости (ее в этом курсе не доказываем).

Предложение 16. *Пусть Z — линейное нормированное пространство, $C \subset Z$ — выпуклое множество, $\operatorname{int} C \neq \varnothing$, $z_0 \notin \operatorname{int} C$. Тогда существует линейный непрерывный функционал $z^* \in Z^* \backslash \{0\}$ такой, что $z^*(z_0) \leq z^*(z)$ для любого $z \in C$.*

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 17. *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*

$$x^*(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j^*(x_j), \quad (x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n. \tag{35}$$

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 (35), x_j^* линеен и $\|x_j^*\| \leq \|x^*\|$.

Обратно, функционал вида (35) линеен и

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

Значит, x^* непрерывен. □

Доказательство теоремы 12. 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

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

Множество C выпукло.

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

$$f_j'(\hat{x})[x_0] < 0,\ 0 \le j \le m, \quad F'(\hat{x})[x_0] = 0. \tag{36}$$

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 $f_j(\hat{x} + tx_0 + r(t)) \overset{(36)}{<} 0$, 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 $\operatorname{int} C \ne \varnothing$. 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 $\mu_k = 1$, $\mu_j = \varepsilon$, $j \ne k$. 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 (34).

Случай 2: $\operatorname{Im} F'(\hat{x}) \ne Y$. 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. $\square$

9 Достаточное условие глобального минимума в задаче с ограничениями типа равенств и неравенств

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

$$\begin{cases} f_0(x) \to \inf, \\ f_j(x) \le 0, \quad 1 \le j \le m, \\ F(x) = 0, \\ x \in U. \end{cases} \tag{37}$$

Следующее утверждение дает достаточное условие минимума в задаче (37).

Предложение 18. *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. $\lambda_0 > 0,\ \lambda_1 \geq 0,\ \ldots,\ \lambda_m \geq 0$*;*

2. $\lambda_j f_j(\hat{x}) = 0,\ 1 \leq j \leq m$*;*

3. $\mathcal{L}(\hat{x}) = \min_{x \in U} \mathcal{L}(x)$*, где* $\mathcal{L}(x) = \sum\limits_{j=0}^{m} \lambda_j f_j(x) + y^*(F(x))$*.*

Тогда $\hat{x}$ *— точка глобального минимума в (37).*

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 $f_0(x) \geq f_0(\hat{x})$. □

Задача 16. (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 $\int\limits_0^\infty x\rho(x)\,dx = C_1$).

10 Задача Лагранжа

10.1 Постановка задачи и формулировка основного результата

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

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

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

при этом отображения

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

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

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

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

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:

$$\begin{cases} \int\limits_{t_0}^{t_1} L_0(t, x(t), u(t))\, dt + l_0(x(t_0), x(t_1)) \to \inf, \\ \int\limits_{t_0}^{t_1} L_k(t, x(t), u(t))\, dt + l_k(x(t_0), x(t_1)) \le 0, \quad 1 \le k \le m', \\ \int\limits_{t_0}^{t_1} L_k(t, x(t), u(t))\, dt + l_k(x(t_0), x(t_1)) = 0, \quad m'+1 \le k \le m, \\ \dot{x}(t) = \varphi(t, x(t), u(t)), \\ x \in C^1([t_0, t_1], \mathbb{R}^n),\ u \in C([t_0, t_1], \mathbb{R}^s). \end{cases} \tag{38}$$

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

$$\mathcal{L}(x,\,u) = \sum_{k=0}^{m} \lambda_k \mathcal{L}_k(x,\,u) + \int\limits_{t_0}^{t_1} p(t)(\dot{x}(t) - \varphi(t,\,x(t),\,u(t)))\,dt. \qquad (39)$$

Выделим интегральное и внеинтегральное слагаемые:

$$\begin{gathered}\mathcal{L}(x,\,u) = \int\limits_{t_0}^{t_1} L(t,\,x,\,\dot{x},\,u)\,dt + l(x(t_0),\,x(t_1)),\\ L(t,\,x,\,\dot{x},\,u) = \sum_{k=0}^{m} \lambda_k L_k(t,\,x,\,u) + p(t)(\dot{x} - \varphi(t,\,x,\,u)), \quad l = \sum_{k=0}^{m} \lambda_k l_k.\end{gathered} \qquad (40)$$

Теорема 13. *Пусть $(\hat{x},\,\hat{u})$ — точка локального минимума в задаче (38). Тогда существует набор $(\lambda_0,\,\dots,\,\lambda_m,\,p(\cdot)) \ne (0,\,\dots,\,0,\,0)$, $\lambda_k \in \mathbb{R}$, $0 \le k \le m$, $p \in C^1([t_0,\,t_1],\,\mathbb{R}^n)$, такой, что для функции $\mathcal{L}$, заданной равенствами (39), (40), 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: $L_u = 0$.*

10.2 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

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

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

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. *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 $(y,\, z) \in \operatorname{Im} C$. □

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.

10.3 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 $E_1,\, \dots,\, E_n,\, \dots \in \Sigma$ 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 $\mu(E) \geq 0$, если $E \subset \Omega_+$, и $\mu(E) \leq 0$, 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.

Для зарядов, как и для σ-аддитивных мер, выполнено свойство непрерывности:

1. Пусть $E_1 \subset E_2 \subset ... \subset E_n \subset E_{n+1} \subset \dots$, $E = \cup_{n\in\mathbb{N}} E_n$, $E_n \in \Sigma$. Тогда $\mu(E) = \lim\limits_{n\to\infty} \mu(E_n)$.

2. Пусть $E_1 \supset E_2 \supset ... \supset E_n \supset E_{n+1} \supset \dots$, $E = \cap_{n\in\mathbb{N}} E_n$, $E_n \in \Sigma$. Тогда $\mu(E) = \lim\limits_{n\to\infty} \mu(E_n)$.

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

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

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 14. *Пусть $x^* \in (C[t_0, t_1])^*$. Тогда на $([t_0, t_1], \mathcal{B})$ существует 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 $\|x\|_C = \max_{t\in[t_0, t_1]} |x(t)|$, где $|x(t)| = \sqrt{x_1^2(t) + \cdots + x_n^2(t)}$. Она эквивалентна норме $\max\{\|x_1\|_C, \ldots, \|x_n\|_C\}$. При этом $C([t_0, t_1], \mathbb{R}^n)$ как множество совпадает с $(C[t_0, t_1])^n$. Значит, если $x^* \in (C([t_0, t_1], \mathbb{R}^n))^*$, то существуют заряды $\mu_1, \ldots, \mu_n$ на $([t_0, t_1], \mathcal{B})$ такие, что

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

10.4 Обобщенная лемма Дюбуа-Реймона

При выводе уравнений Эйлера–Лагранжа для простейшей задачи вариационного исчисления использовалась лемма Дюбуа-Реймона. Здесь мы получим ее обобщение.

Лемма. *Пусть μ — заряд на отрезке $[t_0, t_1]$ со следующим свойством: для любой функции $z \in C^1[t_0, t_1]$ такой, что $z(t_0) = z(t_1) = 0$, выполнено $\int\limits_{[t_0, t_1]} \dot{z}(t) \, d\mu = 0$. Тогда существует число $c \in \mathbb{R}$ такое, что $\mu([t_0, t]) = c(t - t_0)$.*

Доказательство. Сначала покажем, что для любых непересекающихся интервалов $\Delta = (\alpha,\, \beta) \subset [t_0,\, t_1]$ и $\Delta' = (\alpha',\, \beta') \subset [t_0,\, t_1]$ одинаковой длины выполнено $\mu(\Delta) = \mu(\Delta')$.

Пусть без ограничения общности $\beta \leq \alpha'$ и $\mu(\Delta) < \mu(\Delta')$. Положим

$$f_0(t) = \begin{cases} 1, & t \in \Delta, \\ -1, & t \in \Delta', \\ 0, & \text{иначе}. \end{cases}$$

Тогда

$$\int\limits_{[t_0,\, t_1]} f_0\, d\mu = \mu(\Delta) - \mu(\Delta') < 0.$$

Для малых $\delta > 0$ положим

$$f_\delta(t) = \begin{cases} 1, & t \in [\alpha + \delta,\, \beta - \delta], \\ (t-\alpha)/\delta, & t \in (\alpha,\, \alpha + \delta), \\ (\beta - t)/\delta, & t \in (\beta - \delta,\, \beta), \\ -1, & t \in [\alpha' + \delta,\, \beta' - \delta], \\ (\alpha' - t)/\delta, & t \in (\alpha',\, \alpha' + \delta), \\ (t - \beta')/\delta, & t \in (\beta' - \delta,\, \beta'), \\ 0, & \text{иначе}. \end{cases}$$

Тогда f_δ — непрерывная функция и $\int\limits_{t_0}^{t_1} f_\delta(t)\, dt = 0$. При этом $f_0(t)$ совпадает с $f_\delta(t)$ всюду, кроме $E_\delta = (\alpha,\, \alpha+\delta) \cup (\beta-\delta,\, \beta) \cup (\alpha',\, \alpha'+\delta) \cup (\beta'-\delta,\, \beta')$. Значит,

$$\left| \int\limits_{[t_0,\, t_1]} f_0\, d\mu - \int\limits_{[t_0,\, t_1]} f_\delta\, d\mu \right| = \left| \int\limits_{E_\delta} (f_0 - f_\delta)\, d\mu \right| \leq \int\limits_{E_\delta} |f_0 - f_\delta|\, d|\mu| \leq |\mu|(E_\delta).$$

Так как $\cup_{n\in\mathbb{N}} E_{1/n} = \varnothing$ и E_δ друг в друга вложены, то по свойству непрерывности меры $|\mu|(E_\delta) \underset{\delta \to 0}{\to} 0$. Значит,

$$\int\limits_{[t_0,\, t_1]} f_\delta\, d\mu \underset{\delta\to 0}{\to} \int\limits_{[t_0,\, t_1]} f_0\, d\mu$$

и при малых δ выполнено $\int\limits_{[t_0,\, t_1]} f_\delta\, d\mu < 0$. Возьмем такое δ и положим $z(t) = \int\limits_{t_0}^{t} f_\delta(\tau)\, d\tau$. Тогда $z \in C^1[t_0,\, t_1]$, $z(t_0) = z(t_1) = 0$, но $\int\limits_{[t_0,\, t_1]} \dot z(t)\, dt < 0$.

Теперь покажем, что для любого $t \in (t_0,\, t_1)$ выполнено $\mu(\{t\}) = 0$. В самом деле, множество точек t таких, что $\mu(\{t\}) \neq 0$, не более чем счетно. Значит, если $\mu(\{t\}) \neq 0$, то найдется точка $\tau \in (t_0,\, t_1)$ такая, что $\mu(\{\tau\}) = 0$. При малых $\delta > 0$ интервалы $(t - \delta,\, t + \delta)$ и $(\tau - \delta,\, \tau + \delta)$ содержатся в $[t_0,\, t_1]$ и не пересекаются. Значит, $\mu(\tau - \delta,\, \tau + \delta) = \mu(t - \delta,\, t + \delta)$. В силу свойства непрерывности меры, $\mu(\tau - \delta,\, \tau + \delta) \underset{\delta \to 0}{\to} \mu(\{\tau\}) = 0$, $\mu(t - \delta,\, t + \delta) \underset{\delta \to 0}{\to} \mu(\{t\}) \neq 0$.

Покажем, что $\mu(\{t_0\}) = \mu(\{t_1\}) = 0$. Докажем для t_0 (для t_1 доказательство аналогичное). Для малых $\delta > 0$ положим

$$f_\delta(t) = \begin{cases} (t_0 + \delta - t)/\delta, & t \in [t_0,\, t_0 + \delta], \\ 2(t_0 + \delta - t)/\delta, & t \in [t_0 + \delta,\, t_0 + 3\delta/2], \\ 2(t - t_0 - 2\delta)/\delta, & t \in [t_0 + 3\delta/2,\, t_0 + 2\delta], \\ 0, & \text{иначе.} \end{cases}$$

Тогда f_δ непрерывна и $\int\limits_{t_0}^{t_1} f_\delta(t)\, dt = 0$. Далее,

$$\int\limits_{[t_0,\, t_1]} f_\delta\, d\mu = \mu(\{t_0\}) + \int\limits_{(t_0,\, t_0 + 2\delta)} f_\delta\, d\mu,$$

при этом

$$\left| \int\limits_{(t_0,\, t_0 + 2\delta)} f_\delta\, d\mu \right| \le \int\limits_{(t_0,\, t_0 + 2\delta)} |f_\delta|\, d|\mu| \le |\mu|(t_0,\, t_0 + 2\delta) \underset{\delta \to 0}{\to} 0$$

в силу свойства непрерывности меры. Значит, $\int\limits_{[t_0,\, t_1]} f_\delta\, d\mu \underset{\delta \to 0}{\to} \mu(\{t_0\})$.

Положим $z_\delta(t) = \int\limits_{t_0}^{t_1} f_\delta(\tau)\, d\tau$. Тогда $z_\delta \in C^1[t_0,\, t_1]$ и $z_\delta(t_0) = z_\delta(t_1) = 0$. Значит, $\int\limits_{[t_0,\, t_1]} f_\delta\, d\mu = 0$, откуда $\mu(\{t_0\}) = 0$.

Итак, $\mu(\{t\}) = 0$ для любого $t \in [t_0,\, t_1]$ и значения заряда на непересекающихся промежутках одинаковой длины в $[t_0,\, t_1]$ совпадают. Отсюда следует утверждение (сначала доказывается для $t-t_0=q(t_1-t_0)$, $q\in\mathbb{Q}$, затем предельным переходом доказывается для всех $t\in[t_0,\,t_1]$). $\square$

11 Вывод необходимых условий локального минимума в задаче Лагранжа

Напомним, что функция Лагранжа в теореме выглядит следующим образом:

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

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

$$-\dot{p}_k(t)+f_{x_k}(t,\,x,\,u)-\sum_{j=1}^{n}p_j(t)(\varphi_j)_{x_k}(t,\,x,\,u)=0, \tag{41}$$

условия трансверсальности —

$$p_k(t_0)=\frac{\partial l}{\partial x_k(t_0)}(x(t_0),\,x(t_1)),\quad p_k(t_1)=-\frac{\partial l}{\partial x_k(t_1)}(x(t_0),\,x(t_1)), \tag{42}$$

условие стационарности по u —

$$f_{u_k}(t,\,x,\,u)-\sum_{j=1}^{n}p_j(t)(\varphi_j)_{u_k}(t,\,x,\,u)=0. \tag{43}$$

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

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

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

Выпишем производную:

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

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

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

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

$$\int\limits_{t_0}^{t_1} f_{x_k}(t,\, x,\, u)z(t)\, dt + \int\limits_{t_0}^{t_1} \dot{z}(t)\, d\mu_k(t) - \sum_{j=1}^{n} \int\limits_{t_0}^{t_1} (\varphi_j)_{x_k}(t,\, x,\, u)z(t)\, d\mu_j(t) = 0. \tag{44}$$

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

Положим

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

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

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

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

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

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

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

В итоге (44) сводится к равенству

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

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

$$\tilde\mu(E) = -\int\limits_{E} B(t)\, dt + \mu_k(E) - \int\limits_{E} g(t)\, dt.$$

В силу обобщенной леммы Дюбуа-Реймона, существует константа $c \in \mathbb{R}$ такая, что

$$-\int\limits_{[t_0,\, t]} B(\tau)\, d\tau + \mu_k([t_0,\, t]) - \int\limits_{[t_0,\, t]} g(\tau)\, d\tau = c(t - t_0).$$

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 дифференцируема п.в. 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 (41) 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 (41), 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 (42).

Теперь возьмем $h = 0$, $w_j = 0$ при $j \neq k$, $w_k = v$, 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 (43).

12 Задача оптимального управления

12.1 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

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

и

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

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

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

и

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

непрерывны.

Рассмотрим задачу

$$\begin{cases} \int\limits_{t_0}^{t_1} L_0(t,\, x(t),\, u(t))\, dt + l_0(x(t_0),\, x(t_1)) \to \inf, \\ \int\limits_{t_0}^{t_1} L_j(t,\, x(t),\, u(t))\, dt + l_j(x(t_0),\, x(t_1)) \le 0, \quad 1 \le j \le m', \\ \int\limits_{t_0}^{t_1} L_j(t,\, x(t),\, u(t))\, dt + l_j(x(t_0),\, x(t_1)) = 0, \quad m'+1 \le j \le m, \\ \dot{x}(t) = \varphi(t,\, x(t),\, u(t)), \\ u(t) \in U, \quad t \in [t_0,\, t_1], \\ x \in PC^1([t_0,\, t_1],\, \mathbb{R}^n),\ u \in PC([t_0,\, t_1],\, \mathbb{R}^s). \end{cases} \tag{45}$$

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. *Допустимая точка $(\hat x,\,\hat u)$ называется точкой сильного локального минимума в задаче (45), 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 $\lambda_0,\,\ldots,\,\lambda_m\in\mathbb{R}$, $p\in PC^1([t_0,\,t_1],\,\mathbb{R}^n)$ определим функцию Лагранжа

$$\begin{aligned}\mathcal{L}(x,\,u)=\sum_{j=0}^{m}\lambda_j\mathcal{L}_j(x,\,u)+\int\limits_{t_0}^{t_1}p(t)(\dot x(t)-\varphi(t,\,x(t),\,u(t)))\,dt=\\=\int\limits_{t_0}^{t_1}L(t,\,x,\,\dot x,\,u)\,dt+l(x(t_0),\,x(t_1)),\end{aligned}\tag{46}$$

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

Теорема 15. *Пусть $(\hat x,\,\hat u)$ — точка сильного локального минимума в задаче (45). Тогда существуют $\lambda_0,\,\ldots,\,\lambda_m\in\mathbb{R}$, $p\in PC^1([t_0,\,t_1],\,\mathbb{R}^n)$, 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).

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

$$\begin{cases} \int\limits_{t_0}^{t_1} f(t,\, x,\, u)\, dt \to \inf, \\ x(t_0) = x_0, \\ \dot{x} = \varphi(t,\, x,\, u), \\ u(t) \in U, \\ x \in PC^1[t_0,\, t_1],\ u \in PC[t_0,\, t_1]. \end{cases} \tag{47}$$

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

$$\dot{p} = -p(t)\varphi_x(t,\, \hat{x}(t),\, \hat{u}(t)) + f_x(t,\, \hat{x}(t),\, \hat{u}(t)), \tag{48}$$

условие трансверсальности —

$$p(t_0) = \lambda_1,\ p(t_1) = 0, \tag{49}$$

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.

$$f(t,\, \hat{x}(t),\, v) - f(t,\, \hat{x}(t),\, \hat{u}(t)) - p(t)(\varphi(t,\, \hat{x}(t),\, v) - \varphi(t,\, \hat{x}(t),\, \hat{u}(t))) \geq 0, \quad \forall v \in U. \tag{50}$$

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

12.2 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 19. *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*

$$\dot{x} = g(t, x), \quad x(0) = \xi \tag{51}$$

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 20. *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 (51) с $\xi = x_0$ продолжается на весь отрезок $[0, T]$, $R_0 = \|x(\cdot, x_0)\|_{C[0,T]}$. Тогда существует окрестность V точки x_0 такая, что для любого $\xi \in V$ решение задачи (51) продолжается на отрезок $[0, T]$; при этом*

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

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

Доказательство. Пусть $\delta = \delta(3R_0, M(3R_0))$ такое, как в предложении 19. 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. □

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 21. *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 (51) с $\xi = x_0$ продолжается на весь отрезок $[0,\, T]$ (обозначим его $\hat{x}$). Тогда существует окрестность V точки x_0 такая, что для любого $\xi \in V$ решение задачи (51) 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*

$$\dot{y} = g_x(t,\, \hat{x}(t))y, \quad y(0) = 1. \tag{52}$$

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 (51) 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

$$\|x(\cdot,\, x_0 + \lambda) - \hat{x}(\cdot) - \lambda y(\cdot)\|_{C[0,\,T]} = o(\lambda), \quad \lambda \to 0. \tag{53}$$

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:

$$\|x(\cdot,\, x_0+\lambda) - \hat{x} - \lambda y\|_{C[0,\,T]} \le \|x(\cdot,\, x_0+\lambda) - \hat{x} - \lambda y - r(\lambda)\|_{C[0,\,T]} + \|r(\lambda)\|_{C[0,\,T]} =$$
$$= \|x(\cdot,\, x_0+\lambda) - x(\cdot,\, x_0+\lambda+\omega(\lambda))\|_{C[0,\,T]} + \|r(\lambda)\|_{C[0,\,T]} \le C|\omega(\lambda)| + \|r(\lambda)\|_{C[0,\,T]} = \underset{\lambda\to 0}{o}(\lambda).$$

Предложение доказано. □

12.3 Доказательство принцип максимума Понтрягина для задачи со свободным концом

Пусть $(\hat{x},\, \hat{u})$ — точка сильного минимума в задаче (47), $\tau \in (t_0,\, t_1)$ — 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

$$\dot{x}(t) = \varphi(t,\, x(t),\, u_\lambda(t)), \quad x(0) = x_0. \tag{54}$$

Пара $(x_\lambda,\, u_\lambda)$ 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. *Существует $\lambda_0 > 0$ такое, что для любого $\lambda \in (0,\, \lambda_0)$ решение задачи (54) существует на всём отрезке $[t_0,\, t_1]$; при этом*

$$\|x_\lambda - \hat{x}\|_{C[t_0,\, t_1]} \underset{\lambda \to 0+}{\to} 0. \tag{55}$$

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 $\|\hat{x}\|_{C[t_0,\, t_1]}$ 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 20 (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,

$$\left.\frac{d\xi_\lambda}{d\lambda}\right|_{\lambda=0} = \varphi(\tau,\, \hat{x}(\tau),\, v) - \varphi(\tau,\, \hat{x}(\tau),\, \hat{u}(\tau)). \tag{56}$$

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 21, $x'_\xi(\cdot,\, \hat{x}(\tau))$ — решение задачи

$$\dot{h} = \varphi_x(t,\, \hat{x}(t),\, \hat{u}(t))h, \quad h(\tau) = 1. \tag{57}$$

Так как x_λ — решение задачи

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

то

$$\left.\frac{dx_\lambda(t)}{d\lambda}\right|_{\lambda=0} = x'_\xi(\cdot,\, \hat{x}(\tau)) \left.\frac{d\xi_\lambda}{d\lambda}\right|_{\lambda=0} = y(t) \tag{58}$$

— решение задачи

$$\dot{y} = \varphi_x(t,\, \hat{x}(t),\, \hat{u}(t))y, \quad y(\tau) \overset{(56),(57)}{=} \varphi(\tau,\, \hat{x}(\tau),\, v) - \varphi(\tau,\, \hat{x}(\tau),\, \hat{u}(\tau)) \tag{59}$$

(здесь использовалось, что дифференциальное уравнение в (57), (59) линейное).

Теперь перейдем непосредственно к доказательству необходимых условий сильного минимума в (47).

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 (55) 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

$$f_x(t,\, \hat{x}(t),\, \hat{u}(t))\, \frac{dx_\lambda(t)}{d\lambda}\bigg|_{\lambda=0} \overset{(58)}{=} f_x(t,\, \hat{x}(t),\, \hat{u}(t)) y(t),$$

где y — решение (59). Так как $x(\cdot,\, \xi)$ липшицево по ξ и существует $C_2 > 0$ такое, что $|\xi_\lambda - \hat{x}(\tau)| \leq C_2\lambda$, то существует $A > 0$, не зависящее от t и λ, такое, что

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

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,

$$S = f(\tau,\, \hat{x}(\tau),\, v) - f(\tau,\, \hat{x}(\tau),\, \hat{u}(\tau)) + \int\limits_{\tau}^{t_1} f_x(t,\, \hat{x}(t),\, \hat{u}(t))y(t)\, dt \geq 0. \quad (60)$$

Определим функцию p из уравнения Эйлера (48) и условия трансверсальности (49):

$$\dot{p} = -p(t)\varphi_x(t,\, \hat{x}(t),\, \hat{u}(t)) + f_x(t,\, \hat{x}(t),\, \hat{u}(t)), \quad p(t_1) = 0. \quad (61)$$

Тогда

$$\int\limits_{\tau}^{t_1}\frac{d(p(t)y(t))}{dt}dt = -p(\tau)y(\tau) \overset{(59)}{=} -p(\tau)(\varphi(\tau,\, \hat{x}(\tau),\, v) - \varphi(\tau,\, \hat{x}(\tau),\, \hat{u}(\tau))). \quad (62)$$

С другой стороны,

$$\int\limits_{\tau}^{t_1}\frac{d(p(t)y(t))}{dt}dt = \int\limits_{\tau}^{t_1}\dot{p}(t)y(t)\, dt + \int\limits_{\tau}^{t_1} p(t)\dot{y}(t)\, dt \overset{(59),(61)}{=}$$

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

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

Из последнего соотношения и (60), (62) 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.

13 Аэродинамическая задача Ньютона

Требуется найти форму тела вращения с заданной высотой и шириной, образующая которого является монотонно возрастающей функцией, чтобы при движении в воздухе сопротивление было минимальным.

Формализуется задача следующим образом:

$$\begin{cases} \int\limits_0^{T_0} \frac{t}{1+u^2}\,dt \to \inf, \\ x(0)=0,\ x(T_0)=\xi, \\ \dot{x}=u, \\ u \geq 0; \end{cases} \tag{63}$$

здесь $T_0 > 0$, $\xi > 0$ — заданные параметры.

Составим функцию Лагранжа:

$$\mathcal{L} = \lambda_0 \int\limits_0^{T_0} \frac{t}{1+u^2}\,dt + \int\limits_0^{T_0} p(t)(\dot{x}-u)\,dt + \lambda_1 x(0) + \lambda_2 (x(T_0)-\xi).$$

Необходимые условия локального минимума имеют вид $\lambda_0 \geq 0$ (условие неотрицательности), $\dot{p} = 0$ (уравнение Эйлера), $p(0) = \lambda_1$, $p(T_0) = -\lambda_2$ (условие трансверсальности),

$$\min_{v \geq 0} \left(\frac{\lambda_0 t}{1+v^2} - p(t)v \right) = \frac{\lambda_0 t}{1+\hat{u}(t)^2} - p(t)\hat{u}(t)$$

(принцип максимума Понтрягина).

Из уравнения Эйлера получаем, что $p(t) = c$.

Пусть $\lambda_0 = 0$. Тогда $\min_{v \geq 0}(-cv) = -c\hat{u}(t)$. Если $c = 0$, то $p(t) \equiv 0$; из условия трансверсальности следует, что $\lambda_1 = \lambda_2 = 0$, то есть все множители Лагранжа нулевые. Если $c > 0$, то у функции $-cv$ на $[0, +\infty)$ точки минимума нет. Если $c < 0$, то $\hat{u}(t) \equiv 0$. В силу граничного условия в нуле, $\hat{x}(t) \equiv 0$, что противоречит с условием $x(T_0) = \xi > 0$.

Пусть $\lambda_0 > 0$. Без ограничения общности можно взять $\lambda_0 = 1$. Также обозначим $q = -c$. Получаем

$$\min_{v \geq 0} \left(\frac{t}{1+v^2} + qv \right) = \frac{t}{1+\hat{u}(t)^2} + q\hat{u}(t).$$

Для фиксированного $t \in [0, T_0]$ положим $f(v) = \frac{t}{1+v^2} + qv$. Если $q \le 0$, то минимум функции f не достигается, так как она строго убывает. Значит, остается случай $q > 0$.

Найдем участки монотонности функции f на $\mathbb{R}_+$.

Имеем: $f'(v) = -\frac{2tv}{(1+v^2)^2} + q$. Условие $f'(v) = 0$ эквивалентно уравнению

$$q(1+v^2)^2 - 2tv = 0. \tag{64}$$

В левой части стоит строго выпуклая функция, поэтому у нее количество нулей не больше 2. Также заметим, что максимальный корень строго возрастает по t. В самом деле, если $t_1 < t_2$, $q(1+u_1^2)^2 - 2t_1u_1 = 0$, то $q(1+u_1^2)^2 - 2t_2u_1 < 0$. Кроме того, максимальный корень стремится к $+\infty$ при $t \to +\infty$.

Итак, либо f строго возрастает (тогда 0 — точка минимума), либо сначала возрастает, потом убывает и затем снова возрастает. Во втором случае минимум либо при $v = 0$, либо в точке u_*, являющейся максимальным корнем уравнения (64). Сравним значения $f(0)$ и $f(u_*)$. Запишем неравенство $f(0) \le f(u_*)$, получим $\frac{tu_*}{1+u_*^2} \le q$; подставим из (64) $t = \frac{q(1+u_*^2)^2}{2u_*}$ и получим после вычислений $u_* \le 1$.

Итак, если f строго возрастает или $u_* < 1$, то минимум достигается в 0; если $u_* > 1$, то минимум достигается в u_*. Заметим, что $u_* = 1$ при $t = 2q$. Так как u_* строго возрастает по t, то при $t < 2q$ минимум функции f достигается в 0, а при $t > 2q$ — в u_*.

Итак,

$$\hat{u}(t) = \begin{cases} 0, & t < 2q, \\ u_*(t), & t > 2q. \end{cases}$$

В силу условия $x(0) = 0$, при $0 \le t \le 2q$ получаем $\hat{x}(t) = 0$.

При $t \ge 2q$ функцию $x(t)$ запишем параметрически. Выражая t через v из (64), получаем $t(v) = \frac{q}{2}\left(\frac{1}{v} + 2v + v^3\right)$. Далее, $\frac{dx}{dv} = \frac{dx}{dt} \cdot \frac{dt}{dv} = v \cdot \frac{q}{2}\left(-\frac{1}{v^2} + 2 + 3v^2\right) = \frac{q}{2}\left(-\frac{1}{v} + 2v + 3v^3\right)$. Значит, $x(v) = \frac{q}{2}\left(-\ln v + v^2 + \frac{3}{4}v^4\right) + C$; константа C находится из условия $x(1) = 0$ (здесь мы воспользовались тем, что x непрерывна по t и $u_*(2q) = 1$), т.е. $x(v) = \frac{q}{2}\left(-\ln v + v^2 + \frac{3}{4}v^4 - \frac{7}{4}\right)$. Итак,

$$t(v) = \frac{q}{2}\left(\frac{1}{v} + 2v + v^3\right), \quad x(v) = \frac{q}{2}\left(-\ln v + v^2 + \frac{3}{4}v^4 - \frac{7}{4}\right).$$

Теперь покажем, что для любых $T_0 > 0$, $\xi > 0$ найдется $q > 0$ такое, что $x(T_0) = \xi$.

Пусть $a = \frac{\xi}{T_0}$. Возьмем $q = 2$. Покажем, что если $x_0(t) = 0$ при $t \le 2q = 4$, а при $t > 4$ задано параметрически:

$$t(v) = \frac{1}{v} + 2v + v^3, \quad x_0(v) = -\ln v + v^2 + \frac{3}{4}v^4 - \frac{7}{4},$$

то найдется такое $t_* > 4$, что $x_0(t_*) = at_*$. Затем определим q из равенства $\frac{qt_*}{2} = T_0$.

Мы уже говорили, что $\dot{x}_0(t)$ строго возрастает при $t > 4$ и $\dot{x}_0(t) \underset{t\to\infty}{\to} +\infty$. Значит, $x_0(t) - at \underset{t\to+\infty}{\to} +\infty$. Кроме того, $x_0(4) - 4a < 0$. Поэтому уравнение $x_0(t) = at$ имеет корень на $(4, +\infty)$.

Теперь покажем, что найденная экстремаль является точкой минимума. В силу предложения 18, достаточно показать, что она является точкой минимума функции $\mathcal{L}$. Из условий трансверсальности следует, что $\lambda_1 = -q$, $\lambda_2 = q$. Имеем: $\mathcal{L}(x, u) = \mathcal{L}_1(x) + \mathcal{L}_2(u)$, где

$$\mathcal{L}_1(x) = -\int\limits_0^{T_0} q\dot{x}\,dt - qx(0) + q(x(T_0) - \xi) \equiv -q\xi T_0, \quad \mathcal{L}_2(u) = \int\limits_0^{T_0} \left(\frac{t}{1+u^2} + qu\right) dt.$$

Из принципа максимума Понтрягина следует, что $\mathcal{L}_2(u) \ge \mathcal{L}_2(\hat{u})$. Значит, $\mathcal{L}(x, u) \ge \mathcal{L}(\hat{x}, \hat{u})$ для любой допустимой пары (x, u).

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

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

$$\mathcal{L}(x) := \int\limits_{t_0}^{t_1} L(t, x, \dot{x})\,dt \to \inf, \quad x(t_0) = x_0,\ x(t_1) = x_1. \tag{65}$$

Как и раньше, предполагаем, что L, L_x, $L_{\dot{x}}$ непрерывны.

Задачу (65) можно рассматривать как на пространстве $C^1[t_0, t_1]$, так и на пространстве $PC^1[t_0, t_1]$.

Определение. *Допустимая функция* $\hat{x} \in C^1[t_0, t_1]$ *называется точкой слабого локального минимума в задаче (65), если существует такое* $\varepsilon > 0$*, что для любой допустимой функции* $x \in C^1[t_0, t_1]$ *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. *Допустимая функция* $\hat{x} \in PC^1[t_0, t_1]$ *называется точкой сильного минимума в задаче (65), если существует такое* $\varepsilon > 0$*, что для любых допустимых* $x \in PC^1[t_0, t_1]$ *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. Так как $\|x\|_C \leq \|x\|_{C^1}$, то сильный минимум является слабым.

Оказывается, в определении сильного минимума вместо $x \in PC^1[t_0, t_1]$ можно брать $x \in C^1[t_0, t_1]$. Это вытекает из следующей леммы.

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

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

$$M = \max_{t \in [t_0, t_1]} \max\{|x(t)|, |\dot{x}(t)|\}.$$

Для малых $\delta > 0$ построим функцию $y_\delta \in C^1[t_0, t_1]$ 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 $t \in (\tau_j - \delta, \tau_j + \delta)$ выполнено $|x(t) - y_\delta(t)| \leq \int\limits_{\tau_j - \delta}^{t} |\dot{x}(t) - \dot{y}_\delta(t)|\, dt \leq 4M\delta$. Далее,

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

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

(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 $j = 1, \ldots, m$ построить непрерывную функцию $z_{j,\delta} \in C[\tau_j - \delta, \tau_j + \delta]$ так, чтобы $\|z_{j,\delta}\|_C \le M$, $z_{j,\delta}(\tau_j \pm \delta) = z(\tau_j \pm \delta)$, $\int\limits_{\tau_j-\delta}^{\tau_j+\delta} z_{j,\delta}(t)\,dt = \int\limits_{\tau_j-\delta}^{\tau_j+\delta} z(t)\,dt$.

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

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

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

Тогда $\|\xi_{\tau,\sigma}\|_C \le M$, функция $\xi_{\tau,\sigma}$ непрерывна. Для завершения доказательства остается доказать следующее утверждение, которое оставляем в качестве упражнения.

Задача 17. Показать, что если δ достаточно мало, то $\int\limits_{\tau_j-\delta}^{\tau_j+\delta} \xi_{\tau,\tau_j}(t)\,dt > \int\limits_{\tau_j-\delta}^{\tau_j+\delta} z(t)\,dt$, $\int\limits_{\tau_j-\delta}^{\tau_j+\delta} \xi_{\tau_j,\sigma}(t)\,dt < \int\limits_{\tau_j-\delta}^{\tau_j+\delta} z(t)\,dt$ и что найдутся такие $\tau < \tau_j < \sigma$, что $\int\limits_{\tau_j-\delta}^{\tau_j+\delta} \xi_{\tau,\sigma}(t)\,dt = \int\limits_{\tau_j-\delta}^{\tau_j+\delta} z(t)\,dt$. $\square$

Замечание. Для построенной функции y_δ, помимо утверждения леммы, выполнено $\|\dot{y}_\delta\|_C \le \|\dot{x}\|_C$.

15 Необходимые условия сильного минимума

Пусть $L : [t_0, t_1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, L, L_x, $L_{\dot{x}}$ непрерывны.

Определение. *Функция Вейерштрасса* $\mathcal{E} : [t_0, t_1] \times \mathbb{R}^3 \to \mathbb{R}$ *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 16. *Пусть $\hat{x} \in PC^1[t_0,\, t_1]$ — точка сильного минимума в задаче (65). 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 $v \in \mathbb{R}$.*

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 $w \in \mathbb{R}$.

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

$$\lim_{\lambda \to +0} \frac{\mathcal{L}(\hat{x} + h_\lambda) - \mathcal{L}(\hat{x})}{\lambda} \ge 0. \tag{66}$$

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

$$\int\limits_{\tau}^{\sigma} \left(\frac{d}{dt} L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))z(t) + L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))\dot{z}(t) \right) dt =$$

$$= (L_{\dot{x}}(t,\, \hat{x}(t),\, \dot{\hat{x}}(t))z(t))|_{\tau+0}^{\sigma} = -L_{\dot{x}}(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau+0))w.$$

Таким образом, из (66) следует неравенство

$$f(w) := L(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau-0)+w) - L(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau-0)) - L_{\dot{x}}(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau+0))w \geq 0, \quad w \in \mathbb{R} \tag{67}$$

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 (67) следует неравенство $\mathcal{E}(\tau,\, \hat{x}(\tau),\, \dot{\hat{x}}(\tau),\, v) \geq 0$, $v \in \mathbb{R}$, если $\dot{\hat{x}}$ непрерывна в точке τ. □

Задача 18. Сделав замену $\dot{x} = u$, вывести эту теорему из принципа максимума Понтрягина.

Определение. *Пусть $\hat{x} \in PC^1[t_0,\, t_1]$ — 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 $t \in [t_0,\, t_1]$, $v \in \mathbb{R}$.*

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 условия Вейерштрасса означает, что для любого $t \in [t_0, t_1]$ график функции f_t лежит не ниже касательной, проведенной в точке $\dot{\hat{x}}(t)$.

Задача 19. 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 Необходимые и достаточные условия слабого минимума в терминах квадратичных форм

Далее будем дополнительно предполагать, что существуют вторые частные производные $L_{\dot{x}\dot{x}}$, $L_{\dot{x}x}$, L_{xx}, при этом отображения $(t, \xi, \eta) \mapsto L_{\dot{x}\dot{x}}(t, \xi, \eta)$, $(t, \xi, \eta) \mapsto L_{\dot{x}x}(t, \xi, \eta)$, $(t, \xi, \eta) \mapsto L_{xx}(t, \xi, \eta)$ непрерывны.

Пусть $\hat{x}$ — допустимая экстремаль. Будем обозначать $\hat{L}_x(t) = L_x(t, \hat{x}(t), \dot{\hat{x}}(t))$, $\hat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))$, $\hat{L}_{xx}(t) = L_{xx}(t, \hat{x}(t), \dot{\hat{x}}(t))$, $\hat{L}_{\dot{x}x}(t) = L_{\dot{x}x}(t, \hat{x}(t), \dot{\hat{x}}(t))$, $\hat{L}_{\dot{x}\dot{x}}(t) = L_{\dot{x}\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))$.

Пусть $h \in C^1[t_0, t_1]$, $h(0) = h(1) = 0$. 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

$$\mathcal{L}(\hat{x} + \lambda h) - \mathcal{L}(\hat{x}) = \lambda \int\limits_{t_0}^{t_1} (\hat{L}_{\dot{x}}(t)\dot{h}(t) + \hat{L}_x(t)h(t))\, dt +$$

$$+\frac{\lambda^2}{2} \int\limits_{t_0}^{t_1} L_{\dot{x}\dot{x}}(t, \hat{x}(t) + \lambda\theta_{\lambda,t}h(t), \dot{\hat{x}}(t) + \lambda\theta_{\lambda,t}\dot{h}(t))\dot{h}^2(t)\, dt +$$

$$+\frac{\lambda^2}{2} \int\limits_{t_0}^{t_1} 2L_{\dot{x}x}(t, \hat{x}(t) + \lambda\theta_{\lambda,t}h(t), \dot{\hat{x}}(t) + \lambda\theta_{\lambda,t}\dot{h}(t))\dot{h}(t)h(t)\, dt +$$

$$+\frac{\lambda^2}{2} \int\limits_{t_0}^{t_1} L_{xx}(t, \hat{x}(t) + \lambda\theta_{\lambda,t}h(t), \dot{\hat{x}}(t) + \lambda\theta_{\lambda,t}\dot{h}(t))h^2(t)\, dt.$$

Так как $\hat{x}$ — допустимая экстремаль, то $\int\limits_{t_0}^{t_1}(\hat{L}_{\dot{x}}(t)\dot{h}(t) + \hat{L}_x(t)h(t))\,dt = 0$. Далее, отображения $(t,\, \xi,\, \eta) \mapsto L_{\dot{x}\dot{x}}(t,\, \xi,\, \eta)$, $(t,\, \xi,\, \eta) \mapsto L_{\dot{x}x}(t,\, \xi,\, \eta)$, $(t,\, \xi,\, \eta) \mapsto L_{xx}(t,\, \xi,\, \eta)$ равномерно непрерывны на компактах, поэтому для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $|\lambda| < \delta$ для любого $t \in [t_0,\, t_1]$, $h \in C^1[t_0,\, t_1]$, $\|h\|_{C^1} \le 1$, выполнены неравенства

$$\begin{aligned}
&|L_{\dot{x}\dot{x}}(t,\, \hat{x}(t) + \lambda\theta_{\lambda,t}h(t),\, \dot{\hat{x}}(t) + \lambda\theta_{\lambda,t}\dot{h}(t)) - \hat{L}_{\dot{x}\dot{x}}(t)| < \varepsilon,\\
&|L_{\dot{x}x}(t,\, \hat{x}(t) + \lambda\theta_{\lambda,t}h(t),\, \dot{\hat{x}}(t) + \lambda\theta_{\lambda,t}\dot{h}(t)) - \hat{L}_{\dot{x}x}(t)| < \varepsilon,\\
&|L_{xx}(t,\, \hat{x}(t) + \lambda\theta_{\lambda,t}h(t),\, \dot{\hat{x}}(t) + \lambda\theta_{\lambda,t}\dot{h}(t)) - \hat{L}_{xx}(t)| < \varepsilon.
\end{aligned} \tag{68}$$

Предложение 22. *Пусть $\hat{x}$ — точка слабого минимума. Тогда для любого $h \in C^1[t_0,\, t_1]$ такого, что $h(t_0) = h(t_1) = 0$, выполнено*

$$\int\limits_{t_0}^{t_1}(\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\hat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) + \hat{L}_{xx}(t)h^2(t))\,dt \ge 0. \tag{69}$$

Доказательство. Можно считать, что $\|h\|_{C^1} \le 1$. Так как $\hat{x}$ — точка слабого минимума, она удовлетворяет уравнению Эйлера–Лагранжа. Значит, в силу (68),

$$\mathcal{L}(\hat{x}+\lambda h)-\mathcal{L}(\hat{x}) = \frac{\lambda^2}{2}\int\limits_{t_0}^{t_1}(\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t)+2\hat{L}_{\dot{x}x}(t)\dot{h}(t)h(t)+\hat{L}_{xx}(t)h^2(t))\,dt+\underset{\lambda\to 0}{o}(\lambda^2).$$

При малых λ левая часть неотрицательна; отсюда следует (69). $\square$

Теперь докажем достаточное условие слабого минимума.

Предложение 23. *Пусть $\hat{x}$ — допустимая экстремаль. Предположим, что существует $c > 0$ такое, что для любого $h \in C^1[t_0,\, t_1]$ такого, что $h(t_0) = h(t_1) = 0$, выполнено*

$$\int\limits_{t_0}^{t_1}(\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\hat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) + \hat{L}_{xx}(t)h^2(t))\,dt \ge c\int\limits_{t_0}^{t_1}(\dot{h}^2(t) + h^2(t))\,dt. \tag{70}$$

Тогда $\hat{x}$ — точка слабого локального минимума.

Доказательство. Для $\varepsilon > 0$ выберем $\delta > 0$ так, чтобы выполнялось (68). Тогда для любого $h \in C^1[t_0, t_1]$ такого, что $\|h\|_{C^1} \leq 1$, $h(t_0) = h(t_1) = 0$, выполнено

$$\mathcal{L}(\hat{x} + \lambda h) - \mathcal{L}(\hat{x}) \geq$$

$$\geq \frac{\lambda^2}{2} \int_{t_0}^{t_1} (\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\hat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) + \hat{L}_{xx}(t)h^2(t))\,dt - \frac{\varepsilon\lambda^2}{2} \int_{t_0}^{t_1} (\dot{h}^2(t) + 2|\dot{h}(t)h(t)| + h^2(t))\,dt \overset{(70)}{\geq}$$

$$\geq \frac{\lambda^2}{2} \int_{t_0}^{t_1} ((c - \varepsilon)\dot{h}^2(t) - 2\varepsilon|\dot{h}(t)| \cdot |h(t)| + (c - \varepsilon)h^2(t))\,dt.$$

Остается заметить, что при малых $\varepsilon > 0$ квадратичная форма $q(\xi, \eta) = (c - \varepsilon)\xi^2 - 2\varepsilon\xi\eta + (c - \varepsilon)\eta^2$ положительно определена. $\square$

17 Условие Лежандра

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 17. *Пусть $\hat{x}$ — точка слабого минимума. Тогда выполнено условие Лежандра.*

Доказательство. В предыдущем параграфе было показано, что для любого $h \in C^1[t_0, t_1]$ такого, что $h(t_0) = h(t_1) = 0$, выполнено (69).

Пусть условие Лежандра не выполнено, то есть существует точка $\tau \in [t_0, t_1]$ такая, что $\hat{L}_{\dot{x}\dot{x}}(\tau) < 0$. 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 $\hat{L}_{\dot{x}\dot{x}}(t) < -a$, $|\hat{L}_{\dot{x}x}(t)| \leq M$, $|\hat{L}_{xx}(t)| \leq M$. Для малых $\delta > 0$ положим

$$h_\delta(t) = \begin{cases} \delta - |t - \tau|, & t \in [\tau - \delta, \tau + \delta], \\ 0, & \text{иначе.} \end{cases}$$

Тогда

$$\int_{t_0}^{t_1} (\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}_\delta^2(t) + 2\hat{L}_{\dot{x}x}(t)\dot{h}_\delta(t)h_\delta(t) + \hat{L}_{xx}(t)h_\delta^2(t))\,dt = \int_{\tau-\delta}^{\tau+\delta} (\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}_\delta^2(t) + 2\hat{L}_{\dot{x}x}(t)\dot{h}_\delta(t)h_\delta(t) + \hat{L}_{xx}(t)h$$

$$\le -\int\limits_{\tau-\delta}^{\tau+\delta} a\,dt + 2M\int\limits_{\tau-\delta}^{\tau+\delta}(\delta-|t-\tau|)\,dt + 2M\int\limits_{\tau-\delta}^{\tau+\delta}(\delta-|t-\tau|)^2\,dt \le -2a\delta + 4M\delta^2 + 4M\delta^3 < 0,$$

если $\delta > 0$ достаточно мало.

Построенная функция h_δ кусочно-гладкая. Осталось применить лемму о скруглении углов. $\square$

18 Условие Якоби

Обозначим

$$J(h) = \int\limits_{t_0}^{t_1} (\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\hat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) + \hat{L}_{xx}(t)h^2(t))\,dt.$$

Уравнением Якоби называется уравнение Эйлера–Лагранжа для J. Оно имеет вид:

$$-\frac{d}{dt}(\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t) + \hat{L}_{\dot{x}x}(t)h(t)) + \hat{L}_{\dot{x}x}(t)\dot{h}(t) + \hat{L}_{xx}(t)h(t) = 0. \qquad (71)$$

Это линейное уравнение. Сделаем замену $w = \hat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t) + \hat{L}_{\dot{x}x}(t)h(t)$. Тогда (71) записывается в виде системы

$$\dot{w} = \hat{L}_{\dot{x}x}(t)\dot{h}(t) + \hat{L}_{xx}(t)h(t), \quad \hat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t) = w - \hat{L}_{\dot{x}x}(t)h(t).$$

Если $\hat{L}_{\dot{x}\dot{x}}(t) \ne 0$ для любого $t \in [t_0,\, t_1]$, то второе уравнение можно поделить на $\hat{L}_{\dot{x}\dot{x}}(t)$ 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. *Точка $\tau \in (t_0,\, t_1]$ называется сопряженной к t_0, если существует ненулевое решение уравнения (71) такое, что $h(t_0) = h(\tau) = 0$.*

Определение. *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 18. *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 (71), $h_0(t_0) = h_0(\tau) = 0$. Положим $h_*(t) = h_0(t)$ при $0 \leq t \leq \tau$, $h_*(t) = 0$ при $\tau \leq t \leq t_1$.

Покажем, что $J(h_*) = 0$. Так как h_0 — решение уравнения (71), то

$$\int\limits_{t_0}^{\tau} \left(-\frac{d}{dt}(\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}_*(t) + \hat{L}_{\dot{x}x}(t)h_*(t)) + \hat{L}_{\dot{x}x}(t)\dot{h}_*(t) + \hat{L}_{xx}(t)h_*(t) \right) h_*(t)\, dt = 0.$$

Проинтегрировав по частям и учитывая, что $h_*(t_0) = h_*(\tau) = 0$, получаем

$$\int\limits_{t_0}^{\tau} (\hat{L}_{\dot{x}\dot{x}}(t)\dot{h}_*^2(t) + 2\hat{L}_{\dot{x}x}(t)\dot{h}_*(t)h_*(t) + \hat{L}_{xx}(t)h_*^2(t))\, dt = 0.$$

Отсюда следует, что $J(h_*) = 0$.

Пусть $J(h) \geq 0$ для любого $h \in C^1[t_0,\, t_1]$ 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 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. В частности, $\hat{L}_{\dot{x}\dot{x}}(\tau)\dot{h}_*(\tau - 0) = 0$. Так как $\hat{L}_{\dot{x}\dot{x}}(\tau) > 0$, получаем $\dot{h}_0(\tau) = 0$. Но тогда по теореме единственности $h_0 = 0$. Противоречие. □

Задача 20. Показать, что существует такая простейшая задача вариационного исчисления, для которой 0 является точкой глобального минимума, при этом усиленное условие Лежандра не выполнено и условие Якоби не выполнено.

19 Достаточные условия слабого минимума

Мы докажем следующее утверждение.

Теорема 19. *Пусть $\hat{x}$ — допустимая экстремаль в задаче (65). Предположим, что выполнены усиленное условие Лежандра и усиленное условие Якоби. Тогда $\hat{x}$ — точка слабого минимума.*

Обозначим $A(t) = \hat{L}_{\dot{x}\dot{x}}(t)$, $B(t) = \hat{L}_{\dot{x}x}(t)$, $C(t) = \hat{L}_{xx}(t)$.

В силу предложения 23, достаточно показать, что существует $c > 0$ такое, что

$$J(h) := \int\limits_{t_0}^{t_1} (A(t)\dot{h}^2(t) + 2B(t)\dot{h}(t)h(t) + C(t)h^2(t))\,dt \geq c \int\limits_{t_0}^{t_1} (\dot{h}^2(t) + h^2(t))\,dt. \tag{72}$$

Сначала докажем, что левая часть неотрицательна, а потом докажем (72). Будем использовать следующий прием. Пусть $\omega \in C^1[t_0, t_1]$. Тогда

$$\int\limits_{t_0}^{t_1} \frac{d}{dt}(\omega(t)h^2(t))\,dt = \omega h^2|_{t_0}^{t_1} = 0.$$

Значит,

$$J(h) = \int\limits_{t_0}^{t_1} (A(t)\dot{h}^2(t) + 2B(t)\dot{h}(t)h(t) + C(t)h^2(t) - \dot{\omega}(t)h^2(t) - 2\omega(t)\dot{h}(t)h(t))\,dt.$$

Попробуем подобрать функцию ω так, чтобы под интегралом получилось выражение $A(t)\left(\dot{h}(t) + \frac{B(t)-\omega(t)}{A(t)}h(t)\right)^2$. Тогда

$$J(h) = \int\limits_{t_0}^{t_1} A(t)\left(\dot{h}(t) + \frac{B(t)-\omega(t)}{A(t)}h(t)\right)^2 dt \geq 0.$$

Приравняв коэффициенты при h^2, получаем, что ω удовлетворяет дифференциальному уравнению

$$C(t) - \dot{\omega}(t) = \frac{(B(t)-\omega(t))^2}{A(t)}. \tag{73}$$

Если его гладкое решение существует на $[t_0, t_1]$, то проведенные выше рассуждения корректны и $J(h) \geq 0$.

Как построить решение (73)?

Лемма. *Пусть h_0 — решение уравнения Якоби*

$$-\frac{d}{dt}(A(t)\dot{h}(t) + B(t)h(t)) + B(t)\dot{h}(t) + C(t)h(t) = 0 \tag{74}$$

такое, что $h_0(t) > 0$ для любого $t \in [t_0, t_1]$. Пусть ω удовлетворяет равенству

$$\dot{h}_0(t) + \frac{B(t) - \omega(t)}{A(t)} h_0(t) = 0. \tag{75}$$

Тогда $\omega \in C^1[t_0, t_1]$ и выполнено (73).

(Почему ω ищется именно в таком виде? Вспомним, что если h — решение уравнения Якоби, $h(t_0) = h(t_1) = 0$, то $J(h) = 0$; если при этом $J(h) = \int\limits_{t_0}^{t_1} A(t) \left(\dot{h}(t) + \frac{B(t)-\omega(t)}{A(t)} h(t)\right)^2 dt$, то $\dot{h}(t) + \frac{B(t)-\omega(t)}{A(t)} h(t) = 0$.)

Докажем лемму.

Доказательство. Из (75) получаем:

$$A(t)\dot{h}_0(t) + B(t)h_0(t) = \omega(t)h_0(t). \tag{76}$$

В силу (74), левая часть непрерывно дифференцируема; значит, ωh_0 непрерывно дифференцируема. Так как $h_0(t) > 0$ на $[t_0, t_1]$ и $h \in C^1[t_0, t_1]$, то $\omega \in C^1[t_0, t_1]$. Продифференцируем (76) по t:

$$\frac{d}{dt}(A(t)\dot{h}_0(t) + B(t)h_0(t)) = \dot{\omega}(t)h_0(t) + \omega(t)\dot{h}_0(t).$$

Отсюда и из (74) получаем, что $B(t)\dot{h}_0(t)+C(t)h_0(t) = \dot{\omega}(t)h_0(t)+\omega(t)\dot{h}_0(t)$, или

$$C(t) - \dot{\omega}(t) = (\omega(t) - B(t))\frac{\dot{h}_0(t)}{h_0(t)} \overset{(75)}{=} \frac{(\omega(t) - B(t))^2}{A(t)};$$

это и есть уравнение (73). □

Лемма. *Если выполнены усиленное условие Лежандра и усиленное условие Якоби, то существует решение (74), положительное на отрезке $[t_0, t_1]$.*

Доказательство. Пусть h_* — решение (74), $h_*(t_0) = 0$, $\dot h_*(t_0) = 1$. В силу усиленного условия Якоби, $h_*(t) > 0$ для любого $t \in (t_0, t_1]$. Далее, пусть h_{**} — решение (74), $h_{**}(t_0) = 1$, $\dot h_{**}(t_0) = 0$. Тогда существуют $\delta > 0$ и $M > 0$ такие, что $h_{**}(t) \ge \frac{1}{2}$ при $t \in [t_0, t_0+\delta]$, $|h_{**}(t)| \le M$ при $t \in [t_0, t_1]$. Пусть $a = \min_{t\in[t_0+\delta, t_1]} h_*(t)$. Тогда $a > 0$. Положим $h_0 = h_* + \varepsilon h_{**}$, где $\varepsilon > 0$ — достаточно малое число. Тогда $h_0(t) \ge \frac{\varepsilon}{2}$ при $t \in [t_0, t_0+\delta]$, $h_0(t) \ge a - \varepsilon M > 0$ при $t \in [t_0+\delta, t_1]$. □

Теперь докажем (72). Нам нужно проверить, что если $c > 0$ достаточно мало, то

$$\int_{t_0}^{t_1} ((A(t)-c)\dot h^2(t) + 2B(t)\dot h(t)h(t) + (C(t)-c)h^2(t))\,dt \ge 0. \tag{77}$$

Заметим, что при малых c выполнено $A(t) - c > 0$ при $t \in [t_0, t_1]$. В силу доказанного, достаточно проверить, что уравнение

$$-\frac{d}{dt}((A(t)-c)\dot h(t) + B(t)h(t)) + B(t)\dot h(t) + (C(t)-c)h(t) = 0 \tag{78}$$

имеет решение h_c, строго положительное на $[t_0, t_1]$.

Сведем (74) к системе уравнений. Положив $z(t) = A(t)\dot h(t) + B(t)h(t)$, получим $\dot h(t) = -\frac{B(t)}{A(t)}h(t) + \frac{1}{A(t)}z(t)$, $\dot z(t) \overset{(74)}{=} B(t)\dot h(t) + C(t)h(t)$. Значит, (h, z) — решение системы

$$\dot h(t) = -\frac{B(t)}{A(t)}h(t) + \frac{1}{A(t)}z(t), \quad \dot z(t) = \frac{(C(t)A(t) - B^2(t))}{A(t)}h(t) + \frac{B(t)}{A(t)}z(t). \tag{79}$$

Аналогично (78) сводится к системе

$$\begin{aligned} \dot h(t) &= -\frac{B(t)}{A(t)-c}h(t) + \frac{1}{A(t)-c}z(t), \\ \dot z(t) &= \tfrac{((C(t)-c)(A(t)-c) - B^2(t))}{A(t)-c}h(t) + \tfrac{B(t)}{A(t)-c}z(t). \end{aligned} \tag{80}$$

Легко видеть, что коэффициенты матрицы в (80) равномерно сходятся к коэффициентам матрицы в (79) при $c \to 0$.

Пусть h_0 — решение (74), $z_0(t) = A(t)\dot h_0(t) + B(t)h_0(t)$, (h_c, z_c) — решение (80), $h_c(t_0) = h_0(t_0)$, $z_c(t_0) = z_0(t_0)$. Покажем, что $\|h_c - h_0\|_{C[t_0, t_1]} \underset{c\to 0}{\to} 0$. Для этого докажем следующее утверждение.

Лемма. *Пусть $M_\varepsilon(t)$ — матрица $n \times n$ с непрерывными коэффициентами, $t \in [t_0,\, t_1]$, $\varepsilon \geq 0$, $\|M_\varepsilon - M_0\|_C \underset{\varepsilon\to 0}{\to} 0$. Пусть $\dot{x}_\varepsilon(t) = M_\varepsilon(t)x_\varepsilon(t)$, $x_\varepsilon(t_0) = x_0(t_0)$. Тогда $\|x_\varepsilon - x_0\|_C \underset{\varepsilon\to 0}{\to} 0$.*

Доказательство. Если $x_0(t_0) = x_\varepsilon(t_0) = 0$, то по теореме единственности $x_0 = 0$, $x_\varepsilon = 0$. Поэтому далее считаем, что $x_0(t_0) \neq 0$.

Положим $y_\varepsilon = x_\varepsilon - x_0$. Тогда

$$\dot{y}_\varepsilon(t) = M_0(t)y_\varepsilon + (M_\varepsilon(t) - M_0(t))x_\varepsilon(t), \quad y_\varepsilon(t_0) = 0.$$

Покажем, что $\|y_\varepsilon\|_C \underset{\varepsilon\to 0}{\to} 0$.

Сначала покажем, что $\|x_\varepsilon\|_C$ ограничена при малых $\varepsilon > 0$. Пусть

$$w(t) = |x_\varepsilon(t)|^2. \tag{81}$$

Умножим скалярно обе части уравнения $\dot{x}_\varepsilon(t) = M_\varepsilon(t)x_\varepsilon(t)$ на $x_\varepsilon(t)$ и получим $\frac{\dot{w}(t)}{2} = \langle M_\varepsilon(t)x_\varepsilon(t),\, x_\varepsilon(t)\rangle$. Пусть $A_\varepsilon = \sup_{t\in[t_0,\,t_1]} \|M_\varepsilon(t)\|$. Применяя неравенство Коши–Буняковского, получаем при малых $\varepsilon > 0$

$$\frac{\dot{w}(t)}{2} \leq |M_\varepsilon(t)x_\varepsilon| \cdot |x_\varepsilon(t)| \leq (A_0 + 1)w(t).$$

Так как $x_\varepsilon(t_0) \neq 0$, то для любого $t \in [t_0,\, t_1]$ выполнено $w(t) > 0$. Значит, $\frac{\dot{w}(t)}{w(t)} \leq 2(A_0 + 1)$, откуда $\ln w(t) - \ln w(t_0) \leq 2(A_0 + 1)(t - t_0)$. Значит,

$$w(t) \leq w(t_0) \cdot e^{2(A_0+1)(t-t_0)} \leq w(t_0) \cdot e^{2(A_0+1)(t_1-t_0)}.$$

В силу (81), получаем ограниченность $\|x_\varepsilon\|_C$ при малых $\varepsilon > 0$.

Итак, мы получили:

$$\dot{y}_\varepsilon(t) = M_0(t)y_\varepsilon + f_\varepsilon(t), \quad y_\varepsilon(0) = 0,$$

где $\|f_\varepsilon\|_C \underset{\varepsilon\to 0}{\to} 0$. Нам нужно показать, что $\|y_\varepsilon\|_C \underset{\varepsilon\to 0}{\to} 0$.

Пусть $X = \{y \in C^1([t_0,\, t_1], \mathbb{R}^n) : \ y(t_0) = 0\}$. Тогда X банахово (поскольку это замкнутое подпространство в банаховом пространстве $C^1([t_0,\, t_1], \mathbb{R}^n)$). Рассмотрим оператор $T : X \to C([t_0,\, t_1], \mathbb{R}^n)$, $Ty(t) = \dot{y}(t) - M_0(t)y(t)$. Тогда T — линейный непрерывный оператор. По теореме существования и единственности решения дифференциального уравнения, T биективен. По теореме Банаха об обратном операторе, T^{-1} непрерывен. Отсюда следует, что $\|y_\varepsilon\|_C \underset{\varepsilon\to 0}{\to} 0$. □

Итак, мы получили, что $\|h_c-h_0\|_{C[t_0,\,t_1]} \underset{c\to 0}{\to} 0$. Значит, при малых $c>0$ функция h_c будет положительной на $[t_0,\,t_1]$, а это и доказывает теорему.

Задача 21. 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 22. 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 $0\le t\le \delta_\varepsilon$, $\dot{h}_\varepsilon(t)=c_\varepsilon$ при $\delta_\varepsilon\le t\le 3/2$, числа c_ε и δ_ε подобрать. Затем воспользоваться леммой о скруглении углов.)

20 Поле экстремалей и функция действия

Пусть $L:[t_0-\delta,\,t_1+\delta]\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ — непрерывная функция, L_x и $L_{\dot{x}}$ также непрерывны. Как и раньше, рассматриваем задачу

$$\int\limits_{t_0}^{t_1} L(t,\,x(t),\,\dot{x}(t))\,dt \to \inf, \quad x(t_0)=x_{(0)}, \quad x(t_1)=x_{(1)}, \quad x\in C^1[t_0,\,t_1]. \tag{82}$$

Определение. *Пусть $\hat{x}$ — допустимая экстремаль для задачи (82). Скажем, что $\hat{x}$ погружена в поле экстремалей, если существуют окрестность V графика $\hat{x}$ и $\alpha_0\in\mathbb{R}$, $\varepsilon>0$ со следующими свойствами:*

1. *Для каждого $\alpha\in(\alpha_0-\varepsilon,\,\alpha_0+\varepsilon)$ заданы интервал $I_\alpha\subset(t_0-\delta,\,t_1+\delta)$ и экстремаль $x(\cdot,\,\alpha):I_\alpha\to\mathbb{R}$, т.е. решение уравнения Эйлера–Лагранжа*

$$-\frac{d}{dt}L_{\dot{x}}+L_x=0; \tag{83}$$

2. *$x(\cdot,\,\alpha_0)=\hat{x}(\cdot)$;*
3. *отображение $(t,\,\alpha)\mapsto x(t,\,\alpha)$ гладкое;*

4. *график* $x(\cdot,\, \alpha)$ *лежит в* V;

5. *для любого* $(\tau,\, \xi) \in V$ *существует единственное* $\alpha(\tau,\, \xi) \in (\alpha_0 - \varepsilon,\, \alpha_0 + \varepsilon)$ *такое, что* $x(\tau,\, \alpha(\tau,\, \xi)) = \xi$;

6. *отображение* $(\tau,\, \xi) \mapsto \alpha(\tau,\, \xi)$ *гладкое.*

Если $t_0 - \delta \in \overline{I}_\alpha$ для любого $\alpha \in (\alpha_0 - \varepsilon,\, \alpha_0 + \varepsilon)$ и все экстремали $x(\cdot,\, \alpha)$ продолжаются по гладкости в точку $t_0 - \delta$, при этом $x(t_0 - \delta,\, \alpha) = x_*$ для любого α, то поле называется центральным.

Определим функцию наклона поля $u : V \to \mathbb{R}$. 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 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

$$p(\tau,\, \xi) = L_{\dot{x}}(\tau,\, \xi,\, u(\tau,\, \xi)), \quad H(\tau,\, \xi) = u(\tau,\, \xi) L_{\dot{x}}(\tau,\, \xi,\, u(\tau,\, \xi)) - L(\tau,\, \xi,\, u(\tau,\, \xi)) \tag{84}$$

(для запоминания формул: пишем $p = L_{\dot{x}}$, $H = \dot{x} L_{\dot{x}} - L$, затем подставляем $t = \tau$, $x = \xi$, $\dot{x} = u(\tau,\, \xi)$).

Предложение 24. *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, получим

$$\begin{gathered}
\tfrac{\partial}{\partial\tau}\Big|_{t=t_*} x(t,\, \alpha(\tau,\, \xi)) = 0, \quad \tfrac{\partial}{\partial\tau}\Big|_{t=\tau} x(t,\, \alpha(\tau,\, \xi)) = -u(\tau,\, \xi), \\
\tfrac{\partial}{\partial\xi}\Big|_{t=t_*} x(t,\, \alpha(\tau,\, \xi)) = 0, \quad \tfrac{\partial}{\partial\xi}\Big|_{t=\tau} x(t,\, \alpha(\tau,\, \xi)) = 1.
\end{gathered} \tag{85}$$

Далее,

$$\begin{gathered}
S'_\tau(\tau,\, \xi) = L(\tau,\, x(\tau,\, \alpha(\tau,\, \xi)),\, \dot{x}(\tau,\, \alpha(\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\tau}\dot{x}(t,\, \alpha(\tau,\, \xi))+ \\
+L_x(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi))) \frac{\partial}{\partial\tau} x(t,\, \alpha(\tau,\, \xi)) \Big)\, dt \overset{(83)}{=} \\
= L(\tau,\, \xi,\, u(\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 t}\frac{\partial}{\partial\tau} 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\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{(85)}{=} \\
= L(\tau,\, \xi,\, u(\tau,\, \xi)) - L_{\dot{x}}(\tau,\, \xi,\, u(\tau,\, \xi)) u(\tau,\, \xi) = -H(\tau,\, \xi), \\
S'_\xi(\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}\dot{x}(t,\, \alpha(\tau,\, \xi))+ \\
+L_x(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi))) \frac{\partial}{\partial\xi} x(t,\, \alpha(\tau,\, \xi)) \Big)\, dt \overset{(83)}{=} \\
= \int\limits_{t_*}^{\tau} \Big(L_{\dot{x}}(t,\, x(t,\, \alpha(\tau,\, \xi)),\, \dot{x}(t,\, \alpha(\tau,\, \xi))) \frac{\partial}{\partial\xi}\dot{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} 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} x(t,\, \alpha(\tau,\, \xi)) \Big|_{t_*}^{\tau} \overset{(85)}{=}
\end{gathered}$$

$$= L_{\dot x}(\tau,\, \xi,\, u(\tau,\, \xi)) = p(\tau,\, \xi).$$

Равенство доказано. $\square$

Таким образом,

$$dS(\tau,\, \xi) = p(\tau,\, \xi)\, d\xi - H(\tau,\, \xi)\, d\tau. \tag{86}$$

21 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 20. *Пусть $\hat x$ — допустимая экстремаль для задачи (82), окруженная центральным полем экстремалей, $x \in C^1[t_0,\, t_1]$ — допустимая функция, и пусть график x лежит во множестве V из определения 20. Тогда*

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot x(t))\, dt - \int\limits_{t_0}^{t_1} L(t,\, \hat x(t),\, \dot{\hat x}(t))\, dt = \int\limits_{t_0}^{t_1} \mathcal{E}(t,\, x(t),\, u(t,\, x(t)),\, \dot x(t))\, dt. \tag{87}$$

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 (86), $\int\limits_\gamma (p\, dx - H\, dt) = \int\limits_{\hat\gamma} (p\, dx - H\, dt)$. Значит,

$$\int\limits_{t_0}^{t_1} (L_{\dot x}(t,\, x(t),\, u(t,\, x(t)))\dot x(t) - u(t,\, x(t)) L_{\dot x}(t,\, x(t),\, u(t,\, x(t))) + L(t,\, x(t),\, u(t,\, x(t))))\, dt =$$

$$= \int\limits_{t_0}^{t_1} (L_{\dot x}(t,\, \hat x(t),\, u(t,\, \hat x(t)))\dot{\hat x}(t) - u(t,\, \hat x(t)) L_{\dot x}(t,\, \hat x(t),\, u(t,\, \hat x(t))) + L(t,\, \hat x(t),\, u(t,\, \hat x(t))))\, dt,$$

т.е.

$$\int\limits_{t_0}^{t_1} (L_{\dot x}(t,\, x(t),\, u(t,\, x(t)))(\dot x(t) - u(t,\, x(t))) + L(t,\, x(t),\, u(t,\, x(t))))\, dt =$$

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

Вычтем обе части равенства из $\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t)))\, dt$ и получим (87). $\square$

Замечание. При доказательстве достаточного условия слабого минимума мы на самом деле выписывали формулу Вейерштрасса для $J(h)$:

$$\int\limits_{t_0}^{t_1} (A(t)\dot{h}^2(t) + 2B(t)h(t)\dot{h}(t) + C(t)h^2(t))\, dt = \int\limits_{t_0}^{t_1} A(t) \left(\dot{h}(t) + \frac{B(t)-\omega(t)}{A(t)} h(t)\right)^2 dt.$$

В самом деле, функция Вейерштрасса равна

$$\mathcal{E}(t,\, x,\, u,\, v) =$$

$$A(t)v^2 + 2B(t)xv - A(t)u^2 - 2B(t)xu - 2A(t)u(v-u) - 2B(t)x(v-u) =$$
$$= A(t)(v-u)^2.$$

Пусть h_1 — решение уравнения Якоби, $h_1(t) > 0$ для любого $t \in [t_0,\, t_1]$, функция ω определена равенством $\dot{h}_1(t) + \frac{B(t)-\omega(t)}{A(t)} h_1(t) = 0$. Положим $h(t,\, \alpha) = \alpha h_1(t)$. Тогда функция наклона поля имеет вид $u(\tau,\, \xi) = \xi \frac{\dot{h}_1(\tau)}{h_1(\tau)} = -\frac{B(\tau)-\omega(\tau)}{A(\tau)} \xi$. Значит,

$$\int\limits_{t_0}^{t_1} \mathcal{E}(t,\, h(t),\, u(t,\, h(t)),\, \dot{h}(t))\, dt = \int\limits_{t_0}^{t_1} A(t) \left(\dot{h}(t) + \frac{B(t)-\omega(t)}{A(t)} h(t)\right)^2 dt.$$

Определение. *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 $x \in \mathbb{R}$, $u \in \mathbb{R}$ таких, что $|x - \hat{x}(t)| < \varepsilon$, $|u - \dot{\hat{x}}(t)| < \varepsilon$, и для любого $v \in \mathbb{R}$ выполнено $\mathcal{E}(t,\, x,\, u,\, v) \geq 0$.*

Теорема 21. *Пусть $\hat{x}$ — допустимая экстремаль, погруженная в 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 20. Значит, выполнена формула Вейерштрасса. Кроме того, если $\|x-\hat{x}\|_C$ мало, то $|x(t)-\hat{x}(t)| < \varepsilon$, $|u(t, x(t)) - \dot{\hat{x}}(t)| < \varepsilon$. 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

$$\int\limits_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt - \int\limits_{t_0}^{t_1} L(t, \hat{x}(t), \dot{\hat{x}}(t))\, dt = \int\limits_{t_0}^{t_1} \mathcal{E}(t, x(t), u(t, x(t)), \dot{x}(t))\, dt \geq 0.$$

□

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, проведенной в точке u.

Замечание. Если L выпукла по $\dot{x}$, то $\mathcal{E}(t, x, u, v) \geq 0$ для любых x, u, $v \in \mathbb{R}$, $t \in [t_0, t_1]$. Пусть также в качестве V из определения 20 можно взять $(t_0-\delta, t_1+\delta)\times\mathbb{R}$. Тогда в рассуждениях выше допустимая функция x может быть произвольной (а не из окрестности $\hat{x}$ относительно нормы C). Значит, $\hat{x}$ — точка глобального минимума.

Следующую теорему приводим без доказательства.

Теорема 22. *Пусть $L \in C^3([t_0 - \delta, t_1 + \delta] \times \mathbb{R} \times \mathbb{R})$. Пусть $\hat{x}$ — допустимая экстремаль в задаче (82). Предположим, что выполнены усиленное условие Лежандра и усиленное условие Якоби. Тогда $\hat{x}$ можно погрузить в центральное поле экстремалей.*

Отсюда получаем достаточное условие сильного минимума.

Теорема 23. *Пусть $L \in C^3([t_0-\delta, t_1+\delta]\times\mathbb{R}\times\mathbb{R})$. Пусть $\hat{x}$ — допустимая экстремаль в задаче (82). Предположим, что выполнены усиленное условие Лежандра, усиленное условие Якоби и усиленное условие Вейерштрасса. Тогда $\hat{x}$ — точка сильного минимума.*

Задача 23. Рассмотрим задачу $\int\limits_0^1 (\dot{x}^2 - x\dot{x}^3)\, dt \to \text{extr}$, $x(0) = x(1) = 0$. 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.

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.

22 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:
$$\int\limits_0^\infty \dot{x}^2\,dt \geq \int\limits_0^\infty \frac{x^2}{4t^2}\,dt.$$

Решать эту задачу можно следующим образом. Рассмотрим функционал
$$\mathcal{L}(x) = \int\limits_0^\infty \left(\dot{x}^2 - \frac{x^2}{4t^2}\right)\,dt$$
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
$$\int\limits_0^\infty \left(\dot{x}^2 - \frac{x^2}{4t^2}\right)\,dt = \int\limits_0^\infty \left(\dot{x} - \frac{x}{2t}\right)^2\,dt,$$
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:
$$\int\limits_0^\infty \left(\dot{x} - \frac{x}{2t}\right)^2\,dt = \int\limits_0^\infty \left(\dot{x}^2 - \frac{1}{2t}\cdot 2x\dot{x} + \frac{x^2}{4t^2}\right)\,dt =$$
$$= \int\limits_0^\infty \left(\dot{x}^2 + \frac{x^2}{4t^2}\right)\,dt - \int\limits_0^\infty \frac{x^2}{2t^2}\,dt + \left.\frac{x^2}{2t}\right|_0^\infty = \int\limits_0^\infty \left(\dot{x}^2 - \frac{x^2}{4t^2}\right)\,dt + \left.\frac{x^2}{2t}\right|_0^\infty.$$

Остается проверить, что $\frac{x^2}{2t}\Big|_0^\infty = 0$. 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 $t \to +\infty$.

23 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

$$\dot{\hat{x}}(t)L_{\dot{x}}(\hat{x}(t), \dot{\hat{x}}(t)) - L(\hat{x}(t), \dot{\hat{x}}(t)) = \mathrm{const}. \tag{88}$$

В самом деле, продифференцируем левую часть (88) 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 (88) называется законом сохранения энергии.

Замечание. Выражения $p = L_{\dot{x}}$ и $H = \dot{x}L_{\dot{x}} - L$ у нас уже встречались при вычислении производных от функции действия: $dS = p\,dx - H\,dt$.

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

Предложение 25. *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 $\hat{x} \in C^2([t_0, t_1], \mathbb{R}^n)$. $\square$

Предложение 26. *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. (задача о геодезических на плоскости Лобачевского.) Найти допустимые экстремали в задаче

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

Решение. Имеем $L_{\dot x} = \frac{\dot x}{x\sqrt{1+\dot x^2}}$, $L_{\dot x\dot x} = \frac{1}{x(1+\dot x^2)^{3/2}} > 0$. Значит, $\hat x \in C^2[t_0,\, t_1]$ и $\dot{\hat x}(t)L_{\dot x}(\hat x(t),\, \dot{\hat x}(t)) - L(\hat x(t),\, \dot{\hat x}(t)) = \text{const}$.

Проверим, что допустимая экстремаль не может обращаться в константу ни на каком невырожденном интервале. Это видно из уравнения Эйлера:

$$-\frac{d}{dt}\frac{\dot x}{x\sqrt{1+\dot x^2}} - \frac{\sqrt{1+\dot x^2}}{x^2} = 0$$

(тогда бы получилось равенство $1/\hat x^2(t) \equiv 0$).

Из уравнения $\dot x L_{\dot x} - L = \text{const}$ получаем

$$\dot x \cdot \frac{\dot x}{x\sqrt{1+\dot x^2}} - \frac{\sqrt{1+\dot x^2}}{x} = \text{const}.$$

Значит, $\frac{1}{x\sqrt{1+\dot x^2}} = \text{const}$. Получаем $1+\dot x^2 = \frac{c^2}{x^2}$, или $\dot x = \pm\sqrt{1-\frac{c^2}{x^2}}$. Решаем это дифференциальное уравнение и получаем

$$t - a = \pm\int \frac{dx}{\sqrt{1-\frac{c^2}{x^2}}} = \pm\frac{1}{2}\int\frac{dx^2}{\sqrt{x^2-c^2}} = \pm\sqrt{x^2-c^2}.$$

Возводим в квадрат и получаем $x^2 + (t-a)^2 = c^2$. То есть геодезические — дуги окружностей с центром на горизонтальной оси.

Утверждается, что найденная допустимая экстремаль будет точкой глобального минимума. В самом деле, $L_{\dot x\dot x} > 0$ при $x > 0$, так что L выпукла по $\dot x$.

Пусть $\hat x(t) = \sqrt{c^2-(t-a)^2}$, $t_* < t_0$, $c^2-(t_*-a)^2 > 0$, $x_* = \sqrt{c^2-(t_*-a)^2}$, $\delta > 0$ — достаточно малое число. Определим семейство экстремалей (решений уравнения Эйлера) $x(t,\, \alpha)$ таких, что $x(t_*,\, \alpha) = x_*$, $\dot x(t_*,\, \alpha) = \alpha$, $\alpha \in \mathbb{R}$.

Задача 24. Показать, что $\{x(t,\, \alpha) : \ \alpha \in \mathbb{R}\}$ — центральное поле экстремалей, покрывающее область $\{(t,\, x) : \ t_* < t < t_1 + \delta,\, x > 0\}$.

24 Теоремы существования точки минимума

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 $f(x) \le \varliminf_{n\to\infty} f(x_n)$.

Скажем, что T секвенциально компактно, если для любой последовательности $\{x_n\}_{n\in\mathbb{N}} \subset T$ существует сходящаяся подпоследовательность.

Предложение 27. *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}. Положим $\hat{x} = \lim\limits_{k\to\infty} x_{n_k}$. Так как f секвенциально полунепрерывна снизу, то $f(\hat{x}) \le \varliminf\limits_{k\to\infty} f(x_{n_k}) = \inf_{x\in T} f(x)$. Значит, $f(\hat{x}) = \inf_{x\in T} f(x)$. $\square$

Пространства функций бесконечномерные, поэтому шары в них некомпактны относительно топологии, заданной нормой. Но в некоторых случаях замкнутые шары слабо секвенциально компактны. Примером пространств с таким свойством является гильбертово пространство.

Пусть X — нормированное пространство. Напомним, что x_n слабо сходится к x (обозначение: $x_n \underset{n\to\infty}{\overset{w}{\to}} x$), если для любого $f \in X^*$ выполнено $f(x_n) \underset{n\to\infty}{\to} f(x)$. В курсе функционального анализа доказывалось, что слабо сходящаяся последовательность ограничена.

Напомним также теорему Рисса: если H — гильбертово пространство, $f \in H^*$, то существует вектор $y \in H$ такой, что $f(x) = \langle y, x\rangle$, $x \in H$.

Предложение 28. *Пусть H — гильбертово пространство, $\{x_n\}_{n\in\mathbb{N}} \subset H$, $\|x_n\| \le R$ для любого n. Тогда из $\{x_n\}_{n\in\mathbb{N}}$ можно выбрать слабо сходящуюся подпоследовательность $x_{n_k} \underset{k\to\infty}{\overset{w}{\to}} x$; при этом $\|x\| \le R$.*

Доказательство. Пусть $L = \overline{\operatorname{span}}\,\{x_n\}_{n\in\mathbb{N}}$. Тогда L сепарабельно. Если $\{x_{n_k}\}_{k\in\mathbb{N}}$ слабо сходится в L, то эта подпоследовательность слабо сходится в H.

Если L конечномерно, то из $\{x_n\}_{n\in\mathbb{N}}$ можно даже выбрать сходящуюся по норме подпоследовательность.

Пусть L бесконечномерно. Выберем в L ортонормированный базис $\{e_m\}_{m\in\mathbb{N}}$. Для любого $m \in \mathbb{N}$ числовая последовательность $\langle x_n, e_m\rangle$ ограничена и поэтому из нее можно выбрать сходящуюся подпоследовательность.

Применяем канторовский диагональный процесс. Сначала поочередно выбираем подпоследовательности $\{n_l^{(1)}\}_{l\in\mathbb{N}} \supset \{n_l^{(2)}\}_{l\in\mathbb{N}} \supset \dots \supset \{n_l^{(m)}\}_{l\in\mathbb{N}} \supset$

... такие, что $\{\langle e_m,\, x_{n_l^{(m)}}\rangle\}_{l\in\mathbb{N}}$ сходится. Обозначим $n_l = n_l^{(l)}$. Тогда $\{\langle e_m,\, x_{n_l}\rangle\}_{l\in\mathbb{N}}$ сходится для любого $m\in\mathbb{N}$.

Пусть $\xi_m = \lim\limits_{k\to\infty}\langle e_m,\, x_{n_k}\rangle$. Так как $\sum\limits_{j=1}^{m}|\langle e_j,\, x_{n_k}\rangle|^2 \le R^2$, то предельным переходом получаем, что $\sum\limits_{j\in\mathbb{N}}|\xi_j|^2 \le R^2$. Положим $x = \sum\limits_{m\in\mathbb{N}}\xi_m e_m$. Покажем, что $x_{n_k} \underset{k\to\infty}{\overset{w}{\to}} x$ в L.

По теореме Рисса и по определению слабой сходимости, нужно проверить, что

$$\langle y,\, x_{n_k}\rangle \underset{k\to\infty}{\to} \langle y,\, x\rangle, \quad y\in L. \tag{89}$$

Если $y = e_m$, то это выполнено по построению. Отсюда следует, что (89) верно, если y — конечная линейная комбинация e_m. Пусть, наконец, $y\in L$ — произвольный вектор. Для каждого $\varepsilon > 0$ выберем y_ε — конечную линейную комбинацию e_m, так, чтобы $\|y-y_\varepsilon\| < \varepsilon$. Тогда

$$|\langle y,\, x_{n_k}-x\rangle| \le |\langle y_\varepsilon,\, x_{n_k}-x\rangle| + |\langle y-y_\varepsilon,\, x_{n_k}-x\rangle| \le |\langle y_\varepsilon,\, x_{n_k}-x\rangle| + 2R\varepsilon;$$

первое слагаемое будет меньше ε при достаточно больших k. Отсюда следует (89). $\square$

Следствие 2. *Пусть B — замкнутый шар в гильбертовом пространстве, $f : B\to\mathbb{R}\cup\{+\infty\}$ секвенциально слабо полунепрерывна снизу. Тогда f имеет точку минимума на B.*

Пространством Соболева $\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)$ называется пространство

$$\{x\in AC([t_0,\, t_1],\, \mathbb{R}^n) :\ \dot{x}\in L_2([t_0,\, t_1],\, \mathbb{R}^n),\ x(t_0) = x(t_1) = 0\}$$

с нормой

$$\|x\|_{\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)} = \left(\int\limits_{t_0}^{t_1}\sum_{j=1}^{n}|\dot{x}_j(t)|^2\, dt\right)^{1/2}.$$

Пространство $\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)$ изометрически изоморфно пространству

$$H = \left\{x\in L_2([0,\, 1],\, \mathbb{R}^n) :\ \int\limits_{t_0}^{t_1} x_i(t)\, dt = 0,\ 1\le i\le n\right\}$$

(изоморфизм каждой функции $x \in \mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)$ сопоставляет функцию $\dot{x}$). Так как H — замкнутое подпространство в $L_2([0,\, 1],\, \mathbb{R}^n)$, то это гильбертово пространство. Значит, и $\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)$ гильбертово.

Из определения следует, что $\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)$ содержится в $C([t_0,\, t_1],\, \mathbb{R}^n)$ как множество. Пусть $J : \mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n) \to C([t_0,\, t_1],\, \mathbb{R}^n)$ — оператор вложения, то есть $J(x) = x$.

Предложение 29. *Оператор J компактен.*

Доказательство. Нужно показать, что множество

$$B = \{x \in \mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n) :\ \|x\|_{\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)} \le 1\}$$

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 $k \in \{1,\, \dots,\, n\}$, $t_0 \le t < \tau \le t_1$, $x \in B$ имеем

$$|x_k(\tau) - x_k(t)| = \left| \int\limits_t^\tau \dot{x}_k(s)\, ds \right| \le (\tau - t)^{1/2} \|\dot{x}_k\|_{L_2[t,\, \tau]} \le (\tau - t)^{1/2},$$

откуда следует равностепенная непрерывность. Подставив $t = t_0$ и учитывая, что $x(t_0) = 0$, из этой же оценки получаем равномерную ограниченность. $\square$

Напомним следующий факт из функционального анализа: если оператор компактен, то он переводит слабо сходящуюся последовательность в сходящуюся по норме.

Следствие 3. *Если $x_k \underset{k\to\infty}{\overset{w}{\to}} x$ в $\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)$, то x_k равномерно сходится к x.*

Из неравенства Коши–Буняковского и определения нормы в $\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)$ следует

Предложение 30. *Пусть $x \in L_2([t_0,\, t_1],\, \mathbb{R}^n)$, $f(h) = \int\limits_{t_0}^{t_1} x(t)\dot{h}(t)\, dt$, $h \in \mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)$. Тогда $f \in (\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n))^*$ и*

$$|f(h)| \le \|x\|_{L_2([t_0,\, t_1],\, \mathbb{R}^n)} \|h\|_{\mathring{W}_2^1([t_0,\, t_1],\, \mathbb{R}^n)} \tag{90}$$

(здесь $\|x\|_{L_2([t_0,\, t_1],\, \mathbb{R}^n)}$ — это L_2-норма от $|x(\cdot)| = (x_1^2(\cdot) + \dots + x_n^2(\cdot))^{1/2}$).

25 Существование точки минимума в задаче Дидоны

Сначала рассмотрим задачу:

$$\begin{cases} f_0(x,\, y) := \int\limits_0^1 (x\dot y - y\dot x)\, dt \to \inf, \\ f_1(x,\, y) := \int\limits_0^1 (\dot x^2 + \dot y^2)\, dt \le l^2, \\ (x,\, y) \in \mathring{W}_2^1([0,\, 1],\, \mathbb{R}^2). \end{cases} \tag{91}$$

В силу следствия 2, 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 3). Кроме того, $\{(x_n,\, y_n)\}_{n\in\mathbb{N}}$ ограничена, то есть существует $R > 0$ такое, что

$$\|(x_n,\, y_n)\|_{\mathring{W}_2^1([0,\,1],\,\mathbb{R}^2)} \le R, \quad n \in \mathbb{N}. \tag{92}$$

Получаем:

$$\left| \int\limits_0^1 (x_n\dot y_n - y_n\dot x_n)\, dt - \int\limits_0^1 (x\dot y - y\dot x)\, dt \right| \le$$

$$\le \left| \int\limits_0^1 (x_n - x)\dot y_n\, dt \right| + \left| \int\limits_0^1 x(\dot y_n - \dot y)\, dt \right| + \left| \int\limits_0^1 (y_n - y)\dot x_n\, dt \right| + \left| \int\limits_0^1 y(\dot x_n - \dot x)\, dt \right| \le$$

$$\le \|x_n - x\|_{L_2[0,\,1]} \|\dot y_n\|_{L_2[0,\,1]} + \|y_n - y\|_{L_2[0,\,1]} \|\dot x_n\|_{L_2[0,\,1]} +$$

$$+ \left| \int\limits_0^1 x(\dot y_n - \dot y)\, dt \right| + \left| \int\limits_0^1 y(\dot x_n - \dot x)\, dt \right| \overset{(92)}{\le}$$

$$\le R(\|x_n - x\|_{L_2[0,\,1]} + \|y_n - y\|_{L_2[0,\,1]}) + \left| \int\limits_0^1 x(\dot y_n - \dot y)\, dt \right| + \left| \int\limits_0^1 y(\dot x_n - \dot x)\, dt \right| \underset{n\to\infty}{\to} 0$$

по определению слабой сходимости и предложения 30. Таким образом,

$$\int\limits_0^1 (x_n\dot y_n - y_n\dot x_n)\, dt \underset{n\to\infty}{\to} \int\limits_0^1 (x\dot y - y\dot x)\, dt,$$

поэтому f_0 слабо секвенциально полунепрерывен снизу.

В силу следствия 2, точка минимума в задаче (91) существует. Найдем ее.

Вычислив вариацию по Лагранжу у f_0, f_1, получаем:

$$\begin{gathered} f_0'(x,\,y)[(h,\,w)] = \int\limits_0^1 (h\dot y + x\dot w - w\dot x - y\dot h)\,dt, \\ f_1'(x,\,y)[(h,\,w)] = \int\limits_0^1 (2\dot x\dot h + 2\dot y\dot w)\,dt. \end{gathered} \tag{93}$$

Отсюда и из (90) 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 $\lambda_0 \geq 0$, $\lambda_1 \geq 0$, $\lambda_1(f_1(x,\,y) - l^2) = 0$,

$$\lambda_0 f_0'(x,\,y)[(h,\,w)] + \lambda_1 f_1'(x,\,y)[(h,\,w)] = 0 \tag{94}$$

для любого $(h,\,w) \in \mathring{W}_2^1([0,\,1],\,\mathbb{R}^2)$. Подставив (93) в (94) и проинтегрировав по частям $\int\limits_0^1 h\dot y\,dt$ и $\int\limits_0^1 w\dot x\,dt$ 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 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 (91) и в задаче

$$\int\limits_0^1 (x\dot{y} - y\dot{x})\,dt \to \inf, \quad \dot{x}^2 + \dot{y}^2 = l^2, \quad x(0) = x(1) = y(0) = y(1) = 0, \quad (x,\, y) \in C^1 \tag{95}$$

точка минимума одна и та же. Теперь рассмотрим задачу Дидоны:

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

$$x(0) = x(1) = y(0) = y(1) = 0, \quad (x,\, y) \in C^1, \quad \dot{x}^2 + \dot{y}^2 > 0.$$

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 (95). Следовательно, в ней точка минимума существует и совпадает с точкой минимума в (91), т.е. является окружностью.

26 Теорема Тонелли

Рассмотрим задачу

$$\int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\,dt \to \inf, \quad x \in \mathring{W}_2^1[t_0,\, t_1]. \tag{96}$$

Теорема 24. *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*

$$L(t,\, \xi,\, \eta) \geq C\eta^2 - \beta. \tag{97}$$

Тогда точка минимума в (96) существует.

Доказательство. В силу (97),

$$\int_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\,dt \geq C\int_{t_0}^{t_1} \dot{x}^2(t)\,dt - \beta(t_1 - t_0) =$$

$$= C\|x\|^2_{\mathring{W}_2^1[t_0,\, t_1]} - \beta(t_1 - t_0) \underset{\|x\|_{\mathring{W}_2^1[t_0,\, t_1]}\to\infty}{\longrightarrow} +\infty.$$

Значит, задача (96) при достаточно большом $R > 0$ эквивалентна задаче

$$\int_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\,dt \to \inf, \quad \|x\|_{\mathring{W}_2^1[t_0,\, t_1]} \leq R.$$

Поэтому (см. следствие 2) достаточно доказать, что функционал $\mathcal{L}(x) = \int_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\,dt$ слабо секвенциально полунепрерывен снизу.

Пусть $x_n \underset{n\to\infty}{\overset{w}{\to}} x$ в $\mathring{W}_2^1[t_0,\, t_1]$. Нужно проверить, что

$$\varliminf_{n\to\infty} \int_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\,dt \geq \int_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\,dt. \tag{98}$$

Переходя к подпоследовательностям, можно считать, что существует $\lim\limits_{n\to\infty} \int_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\,dt < \infty$.

В силу следствия 3,

$$\|x_n - x\|_{C[t_0,\, t_1]} \underset{n\to\infty}{\to} 0. \tag{99}$$

В силу (99) и предложения 29, $|x_n(t)|$ ограничено на $[t_0,\, t_1]$ равномерно по n и t.

При доказательстве ключевым будет неравенство

$$L(t,\, \xi,\, \eta) - L(t,\, \xi,\, \eta_0) \geq L_{\dot{x}}(t,\, \xi,\, \eta_0)(\eta - \eta_0), \tag{100}$$

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 $\varepsilon > 0$ оказано неравенство

$$\varliminf_{n\to\infty} \int\limits_{E_\varepsilon} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \ge \int\limits_{E_\varepsilon} L(t,\, x(t),\, \dot{x}(t))\, dt. \tag{101}$$

Выведем отсюда (98). Сначала заметим, что

$$\int\limits_{[t_0,\, t_1]\backslash E_\varepsilon} L(t,\, x_n,\, \dot{x}_n)\, dt \overset{(97)}{\ge} -\beta\gamma_\varepsilon \underset{\varepsilon\to 0}{\to} 0.$$

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 (101) выбираем $N(\varepsilon) \in \mathbb{N}$ такое, что при $n \ge N(\varepsilon)$

$$\int\limits_{E_\varepsilon} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \ge \int\limits_{E_\varepsilon} L(t,\, x(t),\, \dot{x}(t))\, dt - \alpha.$$

Тогда при $\varepsilon \in (0,\, \varepsilon_\alpha]$, $n \ge N(\varepsilon)$

$$\int\limits_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \ge \int\limits_{E_\varepsilon} L(t,\, x(t),\, \dot{x}(t))\, dt - 2\alpha.$$

Значит, при $\varepsilon \in (0,\, \varepsilon_\alpha]$

$$\varliminf_{n\to\infty} \int\limits_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \ge \int\limits_{E_\varepsilon} L(t,\, x(t),\, \dot{x}(t))\, dt - 2\alpha.$$

Отсюда следует, что

$$\lim_{n\to\infty} \int\limits_{t_0}^{t_1} L(t,\, x_n(t),\, \dot{x}_n(t))\, dt \ge \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t))\, dt - 2\alpha.$$

В силу произвольности α, получаем (98).

Теперь докажем (101). 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:

$$\int\limits_{E_\varepsilon} (L(t,\, x_n,\, \dot{x}_n) - L(t,\, x_n,\, \dot{x}))\, dt \overset{(100)}{\geq} \int\limits_{E_\varepsilon} L_{\dot{x}}(t,\, x_n,\, \dot{x})(\dot{x}_n - \dot{x})\, dt =$$

$$= \int\limits_{E_\varepsilon} L_{\dot{x}}(t,\, x,\, \dot{x})(\dot{x}_n - \dot{x})\, dt + \int\limits_{E_\varepsilon} (L_{\dot{x}}(t,\, x_n,\, \dot{x}) - L_{\dot{x}}(t,\, x,\, \dot{x}))(\dot{x}_n - \dot{x})\, dt.$$

Напомним, что x_n слабо сходится к x в $\mathring{W}_2^1[t_0,\, t_1]$. 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 $L_2[t_0,\, t_1]$, $L_{\dot{x}}(t,\, x_n(t),\, \dot{x}(t)) - L_{\dot{x}}(t,\, x(t),\, \dot{x}(t))$ 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,

$$\varliminf_{n\to\infty} \int\limits_{E_\varepsilon} (L(t,\, x_n,\, \dot{x}_n) - L(t,\, x_n,\, \dot{x}))\, dt \geq 0.$$

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_n (так как L непрерывна, то она равномерно непрерывна на компактах).

В итоге получаем (101). $\square$

Пример. Рассмотрим задачу

$$\mathcal{L}(x) := \int\limits_0^1 ((1 - |\dot{x}|^{3/2})^{4/3} + x^2)\, dt \to \inf, \quad x \in \mathring{W}_2^1[0,\,1].$$

Тогда $\mathcal{L}(x) > 0$ для любой функции $x \in \mathring{W}_2^1[0,\,1]$. В самом деле, если $\mathcal{L}(x) = 0$, то $x \equiv 0$ и $(1 - |\dot{x}|^{3/2})^{4/3} \equiv 1$ — противоречие.

Покажем, что $\inf\{J(x) : \ x \in \mathring{W}_2^1[0,\,1]\} = 0$. В самом деле, разобьем $[0,\,1]$ на $2n$ равных отрезков; на отрезках с нечетными номерами положим $\dot{x}_n \equiv 1$, на отрезках с четными номерами полагаем $\dot{x}_n \equiv -1$; $x_n(0) := 0$. Тогда $x_n \in \mathring{W}_2^1[0,\,1]$. При этом $(1 - |\dot{x}_n|^{3/2})^{4/3} \equiv 0$, $x_n(j/n) = 0$, $0 \le j \le n$. Значит, для любого $t \in [0,\,1]$ выполнено $|x_n(t)| \le \frac{1}{2n}$. Поэтому $J(x_n) \underset{n\to\infty}{\to} 0$. То есть J не имеет точки минимума на $\mathring{W}_2^1[0,\,1]$ (здесь нарушилось условие выпуклости по $\dot{x}$).