Вариационное исчисление и оптимальное управление (конспект лекций 2012 г.)

Е. Р. Аваков

Содержание

Глава 1. Аппарат теории экстремальных задач

В этой главе собраны те факты, на которые опираются доказательства утверждений, представленные в этих лекциях.

§ 1.1. Элементы функционального анализа

1.1.1. Нормированные пространства и их сопряженные. Векторное пространство X называется (вещественным) *нормированным пространством*, если каждому элементу $x \in X$ сопоставлено число $\|x\|_X$, называемое нормой x, удовлетворяющее условиям: (a) $\|x\|_X \geq 0$ для любого $x \in X$ и $\|x\|_X = 0$ тогда и только тогда, когда $x = 0$; (b) $\|\alpha x\|_X = |\alpha|\|x\|_X$ для любых $\alpha \in \mathbb{R}$ и $x \in X$; (c) $\|x + y\|_X \leq \|x\|_X + \|y\|_X$ для любых $x, y \in X$ (неравенство треугольника).

Если $x, y \in X$, то величина $\rho(x, y) = \|x - y\|_X$ называется *расстоянием между элементами x и y*.

Пусть $\widehat{x} \in X$ и $r > 0$. Множество $B_X(\widehat{x}, r) = \{\, x \in X \mid \|x - \widehat{x}\|_X < r \,\}$ называется *открытым шаром* в X с центром в $\widehat{x}$ радиуса r.

Пусть $A \subset X$. Точка $x \in A$ называется *внутренней точкой* A, если она входит в A с некоторым открытым шаром с центром в x. Множество внутренних точек множества A называется *внутренностью* A и обозначается $\operatorname{int} A$.

Множество A называется *открытым*, если $\operatorname{int} A = A$ (проверьте, что открытый шар — открытое множество).

Любое открытое множество, содержащее данную точку, называется *окрестностью* этой точки.

Множество $B \subset X$ называется *замкнутым*, если его дополнение открыто.

Нормированное пространство X называется *банаховым пространством*, если каждая фундаментальная последовательность его элементов сходится.

Пусть на векторном пространстве X заданы две нормы $\|\cdot\|_1$ и $\|\cdot\|_2$. Говорят, что эти нормы *эквивалентны*, если существуют такие константы $c_i > 0$, $i = 1, 2$, что $c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1$ для всех $x \in X$.

Множество всех линейных непрерывных функционалов на нормированном пространстве X называется *сопряженным пространством* к X и обозначается X^*. Значение линейного функционала $x^* \in X^*$ на элементе $x \in X$ обозначаем так: $\langle x^*, x\rangle$.

Векторное пространство X^* (относительно естественных операций сложения линейных функционалов и умножения их на число) является банаховым пространством с нормой

$$\|x^*\|_{X^*} = \sup_{\|x\|_X \leq 1} \langle x^*, x\rangle.$$

Пусть X и Y — нормированные пространства. Совокупность всех линейных непрерывных операторов $\Lambda\colon X \to Y$ обозначим через $\mathcal{L}(X,Y)$. Это нормированное пространство с нормой

$$\|\Lambda\| = \sup_{\|x\|_X \le 1} \|\Lambda x\|_Y.$$

Пусть $\Lambda \in \mathcal{L}(X,Y)$. Множества $\operatorname{Im}\Lambda = \{\, y \in Y \mid \exists x \in X : \Lambda x = y \,\}$ и $\operatorname{Ker}\Lambda = \{\, x \in X \mid \Lambda x = 0 \,\}$ называются соответственно *образом* и *ядром* оператора Λ.

Если $\operatorname{Im}\Lambda = Y$, то говорят, что оператор Λ *сюръективен* (или Λ — *эпиморфизм* или Λ — *отображение на*).

Если $\operatorname{Ker}\Lambda = \{0\}$, то говорят, что оператор Λ *инъективен* (или Λ — *взаимно однозначное отображение*).

Заметим, что если оператор $\Lambda \in \mathcal{L}(X,Y)$ сюръективен, то сопряженный оператор инъективен. Действительно, пусть $y^* \in \operatorname{Ker}\Lambda^*$. Для любого $y \in Y$ существует $x \in X$ такое, что $y = \Lambda x$ и мы имеем $\langle y^*, y\rangle = \langle y^*, \Lambda x\rangle = \langle \Lambda^* y^*, x\rangle = 0$, т. е. $y^* = 0$ и значит, $\operatorname{Ker}\Lambda^* = \{0\}$.

Пусть X и Y — нормированные пространства. Говорят, что пространства *изометрически изоморфны*, если существует такое линейное сюръективное отображение $A\colon X \to Y$, что $\|Ax\|_Y = \|x\|_X$. Отсюда следует, что отображение A непрерывно и взаимно однозначно. Изометрически изоморфные пространства, как нормированные пространства, не различимы.

Произведение $X \times Y$ нормированных пространств X и Y (которое, как векторное пространство, есть совокупность пар (x,y), где $x \in X$ и $y \in Y$, с операциями покоординатного сложения и умножения на числа) есть нормированное пространство с нормой $\|(x,y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$. Легко видеть, что если X и Y — банаховы пространства, то $X \times Y$ — банахово пространство.

В пространстве $X \times Y$ иногда рассматривают другие, но эквивалентные введенной, нормы, например, $\sqrt{\|x\|_X^2 + \|y\|_Y^2}$ или $\max(\|x\|_X, \|y\|_Y)$ (проверьте это).

Пусть X^* и Y^* — сопряженные пространства соответственно к X и Y и $\Lambda \in \mathcal{L}(X,Y)$. Отображение $\Lambda^*\colon Y^* \to X^*$, определенное по правилу $\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle$ (т. е. функционалу y^* сопоставляется функционал $x \mapsto \langle y^*, \Lambda x\rangle$), называется *сопряженным оператором к* Λ. При этом, $\Lambda^* \in \mathcal{L}(Y^*, X^*)$.

Рассмотрим пространство $X^* \times Y^*$ с нормой $\|(x^*, y^*)\|_{X^*\times Y^*} = \max(\|x^*\|_{X^*}, \|y^*\|_{Y^*})$. Тогда справедлива

Лемма (о сопряженном к произведению пространств)**.** *Пусть X и Y — нормированные пространства. Отображение, сопоставляющее паре $(x^*, y^*) \in X^* \times Y^*$ функционал $z^* \in (X \times Y)^*$, действующий по правилу*

$$\langle z^*, (x,y)\rangle = \langle x^*, x\rangle + \langle y^*, y\rangle, \qquad (i)$$

есть изометрический изоморфизм пространств $X^* \times Y^*$ *и* $(X \times Y)^*$.

Доказательство. Из (i) легко следует, что $z^* \in (X \times Y)^*$ и $\|z^*\|_{(X\times Y)^*} \le \max(\|x^*\|_{X^*}, \|y^*\|_{Y^*}) = \|(x^*, y^*)\|_{X^*\times Y^*}$. Так как $\langle z^*, (x,0)\rangle = \langle x^*, x\rangle$, то $\|x^*\|_{X^*} \le \|z^*\|_{(X\times Y)^*}$. Аналогично $\|y^*\|_{Y^*} \le \|z^*\|_{(X\times Y)^*}$ и тем самым $\|z^*\|_{(X\times Y)^*} = \|(x^*, y^*)\|_{X^*\times Y^*}$. Итак, отображение изометрично и оно сюръективно, так как если $z^* \in (X\times Y)^*$, то полагая $\langle x^*, x\rangle = \langle z^*, (x,0)\rangle$ и $\langle y^*, x\rangle = \langle z^*, (0,y)\rangle$, получаем, что $x^* \in X^*$, $y^* \in Y^*$ и справедливо (i). □

1.1.2. Примеры банаховых пространств. Приведем здесь примеры банаховых пространств, с которыми, в основном, будем иметь дело в дальнейшем.

1. Пространство $\mathbb{R}^n$. Это совокупность всех упорядоченных наборов $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ из n действительных чисел (если $n = 1$, то это просто множество действительных чисел, и мы пишем $\mathbb{R}$ вместо $\mathbb{R}^1$), называемых *векторами* или *вектор-столбцами*, а числа x_i, $1 \le i \le n$ — *координатами вектора* x. Ради экономии места, элементы $\mathbb{R}^n$ иногда будем записывать так $x = (x_1, \ldots, x_n)^T$, где ^{T} обозначает *транспонирование.* В $\mathbb{R}^n$ естественным образом вводится операция (покоординатного) сложения векторов и операция (покоординатного) умножения вектора на число, превращающие это множество в вещественное векторное пространство.

Пусть $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$. Величина $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$ называется *длиной* или *модулем* или *евклидовой нормой* вектора x. Элементарная проверка показывает, что это, действительно, норма в $\mathbb{R}^n$. Из полноты множества действительных чисел следует, что $\mathbb{R}^n$ — банахово пространство.

Пусть $a = (a_1, \ldots, a_n)$ — вектор-строка из n действительных чисел. Отображение $x \mapsto \langle a, x\rangle = \sum_{i=1}^n a_i x_i$, где $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, есть линейный функционал (и непрерывный, как это сразу следует из неравенства Коши–Буняковского) на $\mathbb{R}^n$. Легко видеть, что и любой линейный функционал l на $\mathbb{R}^n$ задается подобным образом с $a = (l(e_1), \ldots, l(e_n))$, где $e_1 = (1, 0, \ldots, 0)^T, \ldots, e_n = (0, \ldots, 0, 1)^T$ — *стандартный базис* в $\mathbb{R}^n$. Таким образом, пространство $(\mathbb{R}^n)^*$ можно отождествить с множеством наборов из n действительных чисел, но расположенных в строку (с аналогичными операциями сложения и умножения на числа и с той же евклидовой нормой).[1]

[1] Точнее говоря, эти пространства изометрически изоморфны.

2. Пространство $C([t_0,t_1],\mathbb{R}^n)$. Это совокупность всех непрерывных отображений из отрезка $[t_0,t_1]$ в пространство $\mathbb{R}^n$ (непрерывных вектор-функций, как еще говорят) с обычными операциями сложения и умножения на числа и нормой $\|x(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)} = \max_{t\in[t_0,t_1]}|x(t)|$ (если $n=1$, то вместо $C([t_0,t_1],\mathbb{R})$ пишем $C([t_0,t_1])$). Несложная проверка показывает, что $C([t_0,t_1],\mathbb{R}^n)$ — банахово пространство.

Обозначим через $\mathrm{Var}([t_0,t_1],\mathbb{R}^n)$ совокупность вектор-функций ограниченной вариации на $[t_0,t_1]$, т. е. таких функций $v(\cdot)$, для которых конечна величина $V(v(\cdot)) = \sup\sum_{k=1}^m |v(\tau_k)-v(\tau_{k-1})|$, где верхняя грань берется по всем разбиениям отрезка $[t_0,t_1]$ вида $t_0=\tau_0<\tau_1<\ldots<\tau_m=t_1$. Каждая функция $v(\cdot) = (v_1(\cdot),\ldots,v_n(\cdot)) \in \mathrm{Var}([t_0,t_1],\mathbb{R}^n)$ определяет векторную меру на $[t_0,t_1]$ (меру Лебега–Стилтьеса). Интеграл Лебега по этой мере от суммируемой вектор-функции $x(\cdot)=(x_1(\cdot),\ldots,x_n(\cdot))$ записывается так

$$\sum_{k=1}^n\int_{t_0}^{t_1} x_k(t)\,dv_k(t), \quad \text{или короче,} \quad \int_{t_0}^{t_1} x(t)\,dv(t).$$

Пусть $\mathrm{Var}_0([t_0,t_1],\mathbb{R}^n)$ — подмножество $\mathrm{Var}([t_0,t_1],\mathbb{R}^n)$, образованное функциями, непрерывных справа на (t_0,t_1) и равными нулю в t_0. Это банахово пространство с нормой $V(v(\cdot))$.

Теорема (Ф. Рисса)**.** *Отображение, которое сопоставляет каждой функции* $v(\cdot)\in\mathrm{Var}_0([t_0,t_1],\mathbb{R}^n)$ *линейный непрерывный функционал* x^* *на* $C([t_0,t_1],\mathbb{R}^n)$*, действующий по правилу*

$$\langle x^*, x(\cdot)\rangle = \int_{t_0}^{t_1} x(t)\,dv(t), \tag{R}$$

есть изометрический изоморфизм[2] *пространств* $\mathrm{Var}_0([t_0,t_1],\mathbb{R}^n)$ *и* $(C([t_0,t_1],\mathbb{R}^n))^*$.

Будем говорить, что (R) — *каноническое представление* функционала $x^*\in(C([t_0,t_1],\mathbb{R}^n))^*$.

3. Пространство $C^1([t_0,t_1],\mathbb{R}^n)$. Это совокупность всех непрерывно дифференцируемых вектор-функций на отрезке $[t_0,t_1]$ с обычными операциями сложения и умножения на числа и нормой $\|x(\cdot)\|_{C^1([t_0,t_1],\mathbb{R}^n)} = \max(\|x(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)}, \|\dot x(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)})$ (если $n=1$, то вместо $C^1([t_0,t_1],\mathbb{R})$ пишем $C^1([t_0,t_1])$). Снова, простая проверка показывает, что $C^1([t_0,t_1],\mathbb{R}^n)$ — банахово пространство.

Из теоремы Рисса вытекает следующая

[2] В частности представление (R) единственно: $x^*=0 \Leftrightarrow v(\cdot)=0$.

Теорема. *Отображение, которое сопоставляет каждой паре* $(a, v(\cdot)) \in (\mathbb{R}^n)^* \times \mathrm{Var}_0([t_0, t_1], \mathbb{R}^n)$ *линейный непрерывный функционал* x^* *на* $C^1([t_0, t_1], \mathbb{R}^n)$, *действующий по правилу*

$$\langle x^*, x(\cdot)\rangle = \langle a, x(t_0)\rangle + \int_{t_0}^{t_1} \dot{x}(t)\, dv(t), \qquad (R_1)$$

есть изометрический изоморфизм[3] *пространств* $(\mathbb{R}^n)^* \times \mathrm{Var}_0([t_0, t_1], \mathbb{R}^n)$ *и* $(C^1([t_0, t_1], \mathbb{R}^n))^*$.

Здесь также говорим, что (R_1) — *каноническое представление* функционала $x^* \in (C^1([t_0, t_1], \mathbb{R}^n))^*$.

1.1.3. Теорема отделимости. Лемма о нетривиальности аннулятора. Пусть A и B — непустые подмножества нормированного пространства X. Говорят, что ненулевой функционал $x^* \in X^*$ *отделяет множества* A *и* B, если

$$\sup_{x\in A}\langle x^*, x\rangle \le \inf_{x\in B}\langle x^*, x\rangle.$$

Если неравенство строгое, то говорят, что x^* *строго отделяет* A *и* B.

Пусть число $\gamma \in \mathbb{R}$ таково, что $\sup_{x\in A}\langle x^*, x\rangle \le \gamma \le \inf_{x\in B}\langle x^*, x\rangle$. Тогда, геометрически, отделимость множеств A и B означает, что они расположены по разные стороны от гиперплоскости $\{\, x \in X \mid \langle x^*, x\rangle = \gamma \,\}$.

Непустое множество $A \subset X$ называется *выпуклым*, если с любыми двумя своими точками оно содержит и отрезок $[x, y] = \{\, z \in X \mid z = (1-\alpha)x + \alpha y,\ 0 \le \alpha \le 1 \,\}$, соединяющий эти точки.

Теорема (Первая теорема отделимости)**.** *Пусть* A *и* B *— непустые выпуклые подмножества нормированного пространства* X, *причем* $\operatorname{int} A \ne \emptyset$ *и* $(\operatorname{int} A) \cap B = \emptyset$. *Тогда множества* A *и* B *отделимы.*

Отсюда следует

Теорема (Вторая теорема отделимости)**.** *Пусть* A *— непустое замкнутое выпуклое подмножество нормированного пространства* X *и* $\widehat{x} \notin A$. *Тогда множества* A *и* $\widehat{x}$ *строго отделимы.*

Доказательство. Так как A замкнуто, что существует такое $r > 0$, что открытый шар $B_X(\widehat{x}, r)$ не пересекается с A. Тогда по первой теореме отделимости существует ненулевой функционал $x^* \in X^*$ такой, что $\sup_{x\in A}\langle x^*, x\rangle \le \inf_{x\in B_X(\widehat{x},r)}\langle x^*, x\rangle$. Но $\inf_{x\in B_X(\widehat{x},r)}\langle x^*, x\rangle < \langle x^*, \widehat{x}\rangle$, так как ненулевой линейный непрерывный функционал не может достигать минимума во внутренней точке (докажите это). Итак, множества A и $\widehat{x}$ строго отделимы. $\square$

[3] В частности представление (R_1) единственно: $x^* = 0 \Leftrightarrow a = 0$ и $v(\cdot) = 0$.

Пусть L — подпространство нормированного пространства X. Множество

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

называется *аннулятором* L. Легко видеть, что $L^{\perp}$ — замкнутое подпространство в X^*.

Лемма (о нетривиальности аннулятора)**.** *Пусть L — замкнутое подпространство нормированного пространства X, не совпадающее с X. Тогда $L^{\perp} \neq \{0\}$.*

Доказательство. Так как $L \neq X$, то существует $\widehat{x} \notin L$. Множество L, очевидно, выпукло и по условию замкнуто, поэтому по второй теореме отделимости найдется ненулевой функционал $x^* \in X^*$ такой, что

$$\sup_{x\in L}\langle x^*, x\rangle < \langle x^*, \widehat{x}\rangle. \tag{i}$$

Тогда $x^* \in L^{\perp}$. Действительно, если $\langle x^*, x_0\rangle \neq 0$ для некоторого $x_0 \in L$, то так как $\alpha x_0 \in L$ для любого $\alpha \in \mathbb{R}$ мы придем к противоречию с соотношением (i). $\square$

1.1.4. Теорема Банаха об открытом отображении. Лемма о правом обратном. Пусть X и Y — нормированные пространства и $\Lambda \in \mathcal{L}(X, Y)$. Отображение $F\colon X \to Y$ называется *открытым*, если образ любого открытого множества открыт.

Теорема (Банаха об открытом отображении)**.** *Пусть X и Y — банаховы пространства и оператор $\Lambda \in \mathcal{L}(X, Y)$ сюрьективен. Тогда Λ — открытое отображение.*

Следствие (Лемма о правом обратном)**.** *В условиях теоремы Банаха об открытом отображении существуют отображение $R\colon Y \to X$ и константа $\gamma > 0$ такие, что $\Lambda R(y) = y$ и $\|R(y)\|_X \leq \gamma\|y\|_Y$ для любого $y \in Y$.*

Доказательство. По теореме Банаха об открытом отображении множество $\Lambda(B_X(0,1))$ открыто, очевидно, содержит ноль и тем самым содержит некоторый шар $B_Y(0,r)$, $r > 0$, т. е. для каждого $y \in B_Y(0,r)$ найдется элемент $x(y) \in B_X(0,1)$ такой, что $\Lambda x(y) = y$. Положим $R(0) = 0$, а если $y \neq 0$, то $R(y) = (2\|y\|_Y/r)x((r/2\|y\|_Y)y)$. Тогда $\Lambda R(y) = y$ и $\|R(y)\|_X \leq \gamma\|y\|_Y$, где $\gamma = 2/r$. $\square$

Заметим, что отображение R (правый обратный оператор к Λ), вообще говоря, не линейно, не непрерывно, но непрерывно в нуле.

Лемма (о замкнутости образа)**.** *Пусть X и Y — банаховы пространства, $A \in \mathcal{L}(X, Y)$, $B \in \mathcal{L}(X, \mathbb{R}^s)$, $C\colon X \to Y \times \mathbb{R}^s$, $Cx = (Ax, Bx)$ и $\operatorname{Im} A = Y$. Тогда $\operatorname{Im} C$ — замкнутое подпространство в $Y \times \mathbb{R}^s$.*

Доказательство. Пусть $(y,z) \in \operatorname{cl}\operatorname{Im} C$ и пусть $\{x_n\}$ — последовательность в X такая, что $y = \lim_{n\to\infty} Ax_n$ и $z = \lim_{n\to\infty} Bx_n$. Положим $h_n = R(Ax_n - y)$, где R — правый обратный к A, тогда $A(x_n - h_n) = y$. Так как $\|h_n\|_X = \|R(Ax_n - y)\|_X \le \gamma\|Ax_n - y\|_Y$, то $h_n \to 0$ при $n \to \infty$ и поэтому $\lim_{n\to\infty} B(x_n - h_n) = z$. Таким образом, z принадлежит замыканию образа множества $\{\, x \in X \mid Ax = y \,\}$ при отображении B, которое есть линейное многообразие в $\mathbb{R}^s$ и тем самым замкнуто. Следовательно, существует такое $\overline{x} \in X$, что $A\overline{x} = y$ и $z = B\overline{x}$, т. е. $(y,z) \in \operatorname{Im} C$. $\square$

Лемма (об аннуляторе ядра эпиморфизма)**.** *Пусть X и Y — банаховы пространства, $A \in \mathcal{L}(X,Y)$ и $\operatorname{Im} A = Y$. Тогда $(\operatorname{Ker} A)^\perp = \operatorname{Im} A^*$.*

Доказательство. Включение $\operatorname{Im} A^* \subset (\operatorname{Ker} A)^\perp$ проверяется без труда. Пусть $x^* \in (\operatorname{Ker} A)^\perp$. Образ оператора $M\colon X \to \mathbb{R} \times Y$, $Mx = (\langle x^*, x\rangle, Ax)$, замкнут по лемме о замкнутости образа и не совпадает с $\mathbb{R} \times Y$, так как $(1, 0_Y) \notin \operatorname{Im} A$. Следовательно, по лемме о нетривиальности аннулятора (см. п. 1.1.3) существует ненулевой функционал $(\alpha, y^*) \in \mathbb{R} \times Y^*$ такой, что $\alpha\langle x^*, x\rangle + \langle y^*, Ax\rangle = 0$ для всех $x \in X$. При этом $\alpha \ne 0$, ибо в противном случае функционал y^* был бы нулевым в силу сюръективности A. Таким образом, $x^* = A^*(-\alpha^{-1}y^*) \in \operatorname{Im} A^*$. $\square$

§ 1.2. Дифференциальное исчисление в нормированных пространствах

1.2.1. Основные понятия и теоремы. Пусть X, Y — нормированные пространства, U — открытое подмножество X и задано отображение $F\colon U \to Y$.

Определение. *Отображение $F\colon U \to Y$ называется дифференцируемым в точке $\widehat{x} \in U$, если найдется такой оператор $\Lambda \in \mathcal{L}(X,Y)$, что для всех $h \in X$, для которых $x + h \in U$ справедливо представление*

$$F(\widehat{x} + h) = F(\widehat{x}) + \Lambda h + r(h),$$

где $r(h) = o(\|h\|_X)$ ($\|r(h)\|_Y/\|h\|_X \to 0$ при $h \to 0$). Оператор Λ определен этим равенством однозначно. Он называется производной отображения F в точке $\widehat{x}$ и обозначается $F'(\widehat{x})$.

Если отображение F дифференцируемо в каждой точке U, то определено отображение $F'\colon U \to \mathcal{L}(X,Y)$, сопоставляющее $x \in U$ производную $F'(x)$. Если это отображение непрерывно в $\widehat{x} \in U$ (на U), то говорят, что отображение F *непрерывно дифференцируемо в* $\widehat{x}$ (на U).

Определение. *Отображение $F\colon U\to Y$ называется строго дифференцируемым в точке $\widehat{x}\in U$, если найдется такой оператор $\Lambda\in\mathcal{L}(X,Y)$, что для любого $\varepsilon>0$ существует $\delta=\delta(\varepsilon)>0$, обладающее тем свойством, что если $\|x_i-\widehat{x}\|_X<\delta$, $i=1,2$, то справедливо неравенство*

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

Отсюда следует (полагая $x_2=\widehat{x}$), что F дифференцируемо в $\widehat{x}$ и тем самым $\Lambda=F'(\widehat{x})$.

Обратное не верно. Действительно, из определения строгой дифференцируемости в $\widehat{x}$ легко следует, что отображение непрерывно в некоторой окрестности $\widehat{x}$. Рассмотрим функцию $f\colon\mathbb{R}\to\mathbb{R}$, $f(x)=x^2D(x)$, где $D(\cdot)$ — функция Дирихле (равная единице, если x рационально и нулю, если иррационально). Легко проверить, что f дифференцируема в нуле и, очевидно, разрывна во всех остальных точках и значит, не может быть строго дифференцируемой в нуле.

Если отображение непрерывно дифференцируемо в $\widehat{x}$, то оно строго дифференцируемо в этой точке (см. следствие ниже).

Теорема (о среднем)**.** *Пусть X,Y — нормированные пространства, U — открытое подмножество X и отображение $F\colon U\to Y$ дифференцируемо в каждой точке отрезка $[x_1,x_2]\subset U$. Тогда*

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

Следствие. *Пусть X,Y — нормированные пространства, U — открытое подмножество X и отображение $F\colon U\to Y$ непрерывно дифференцируемо в точке $\widehat{x}\in U$. Тогда F строго дифференцируемо в $\widehat{x}$.*

Доказательство. Пусть $\varepsilon>0$ и $\delta=\delta(\varepsilon)>0$ такое, что $\|F'(x)-F'(\widehat{x})\|<\varepsilon$ для $x\in B_X(\widehat{x},\delta)$. Если $x_i\in B_X(\widehat{x},\delta)$, $i=1,2$, то легко видеть, что $[x_1,x_2]\subset B_X(\widehat{x},\delta)$ и тогда применяя теорему о среднем к отображению $x\mapsto F(x)-F'(\widehat{x})x$, получаем, что

$$\begin{aligned}&\|F(x_1)-F(x_2)-F'(\widehat{x})(x_1-x_2)\|_Y\le\\&\qquad\le\sup_{x\in[x_1,x_2]}\|F'(x)-F'(\widehat{x})\|\|x_1-x_2\|_X\le\varepsilon\|x_1-x_2\|_X,\end{aligned}$$

т. е. F строго дифференцируемо в $\widehat{x}$. □

Дадим определение второй производной. Пусть отображение $F'\colon U\to\mathcal{L}(X,Y)$ дифференцируемо в точке $\widehat{x}$. Тогда говорят, что F дважды дифференцируема в $\widehat{x}$ и соответствующую (вторую) производную обозначают $F''(\widehat{x})$. Ясно, что $F''(\widehat{x})\in\mathcal{L}(X,\mathcal{L}(X,Y))$.

Пространство $\mathcal{L}(X,\mathcal{L}(X,Y))$ изометрически изоморфно пространству $\mathcal{L}^2(X,Y)$ всех непрерывных билинейных отображений $B\colon X\times$

$X \to Y$ с нормой $\|B\| = \sup_{\|x_1\|_X \le 1, \|x_2\|_X \le 1} \|B[x_1, x_2]\|_Y$. Действительно, пусть отображение A сопоставляет $\Lambda \in \mathcal{L}(X, \mathcal{L}(X, Y))$ отображение $B\colon X \times X \to Y$, действующее по правилу: $B[x_1, x_2] = \Lambda x_1[x_2]$ (действие оператора Λx_1 на элементе x_2). Очевидно, что B — билинейное отображение. Далее, $\|\Lambda\| = \sup_{\|x_1\|_X} \|\Lambda x_1\|_{\mathcal{L}(X,Y)} = \sup_{\|x_1\|_X, \|x_2\|_X \le 1} \|\Lambda x_1[x_2]\|_Y = \|B\|$, т. е. A изометрично отображает $\mathcal{L}(X, \mathcal{L}(X, Y))$ в $\mathcal{L}^2(X, Y)$. Линейность и сюръективность A проверяются без труда.

Теорема (Формула Тейлора). *Пусть X и Y — нормированные пространства, U — окрестность точки $\widehat{x} \in X$. Если отображение $F\colon U \to Y$ дважды дифференцируемо в точке $\widehat{x}$, то $F''(\widehat{x})$ — симметричное билинейное отображение и справедлива формула Тейлора*

$$F(\widehat{x} + h) = F(\widehat{x}) + F'(\widehat{x})h + \frac{1}{2}F''(\widehat{x})[h, h] + r(h),$$

где $r(h) = o(\|h\|_X^2)$, т. е. $\|r(h)\|_Y / \|h\|_X^2 \to 0$ при $h \to 0$.

Теорема (о суперпозиции дифференцируемых отображений). *Пусть X, Y и Z — нормированные пространства, U — окрестность точки $\widehat{x} \in X$, $F\colon U \to Y$, V — окрестность точки $F(\widehat{x})$, $G\colon V \to Z$ и $H = G \circ F\colon U \to Z$. Тогда, если F дифференцируемо (строго дифференцируемо) в точке $\widehat{x}$, G дифференцируемо (строго дифференцируемо) в точке $F(\widehat{x})$, то отображение H дифференцируемо (строго дифференцируемо) в точке $\widehat{x}$ и $H'(\widehat{x}) = G'(F(\widehat{x})) \circ F'(\widehat{x})$.*

Пусть X, Y и Z — нормированные пространства, W — окрестность точки $(\widehat{x}, \widehat{y}) \in X \times Y$ и $F\colon W \to Z$. Если отображение $x \to F(x, \widehat{y})$ (определенное на проекции W на X) дифференцируемо в точке $\widehat{x}$, то соответствующую производную называют *частной производной отображения F по x в точке $(\widehat{x}, \widehat{y})$* и обозначают $F_x(\widehat{x}, \widehat{y})$. Аналогично, частную производную F по y в точке $(\widehat{x}, \widehat{y})$ обозначаем $F_y(\widehat{x}, \widehat{y})$.

Теорема (о полном дифференциале). *Пусть X, Y и Z — нормированные пространства и W — открытое подмножество $X \times Y$. Отображение $F\colon W \to Z$ непрерывно дифференцируемо на W тогда и только тогда, когда его частные производные F_x и F_y непрерывно дифференцируемы на W и при этом $F'(x, y)[\xi, \eta] = F_x(x, y)[\xi] + F_y(x, y)[\eta]$ для любых $(x, y) \in W$, $\xi \in X$ и $\eta \in Y$.*

1.2.2. Дифференцируемость некоторых отображений. Рассмотрим здесь три примера дифференцируемых отображений.

Пример 1. *Отображение конечномерных пространств.* Пусть U — открытое подмножество $\mathbb{R}^n$ и $F\colon U \to \mathbb{R}^m$. Оператор $\Lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ можно отождествить с его матрицей (размера $m \times n$) в стандартных базисах $\mathbb{R}^n$ и $\mathbb{R}^m$. В этом случае Λx — это произведение матрицы Λ на вектор-столбец x.

Пусть заданы функции $f_i\colon U \to \mathbb{R}$, $i = 1, \dots, m$. Тогда определено отображение $F\colon U \to \mathbb{R}^m$ по формуле $F(x) = (f_1(x), \dots, f_m(x))^T$, $x \in U$. Если это отображение дифференцируемо в точке $\widehat{x} \in U$, то $F'(\widehat{x})$ — матрица, строки которой суть векторы $f_1'(\widehat{x}), \dots, f_k'(\widehat{x})$. Эту матрицу называют *матрицей Якоби* отображения F в точке $\widehat{x}$.

Если функция $f\colon U \to \mathbb{R}$ дважды дифференцируемо в $\widehat{x}$, то вторая производная — билинейная форма с матрицей $(\partial^2 f(x)/\partial x_j \partial x_i)$, $i, j = 1, \dots, n$, которую называют *матрицей Гесса* или *гессианом функции f в точке $\widehat{x}$*.

Пример 2. *Аффинное отображение.* Пусть X и Y — нормированные пространства $\Lambda \in \mathcal{L}(X, Y)$, $a \in Y$. Отображение $\mathcal{A}\colon X \to Y$, действующее по правилу $\mathcal{A}(x) = \Lambda x + a$, называется аффинным отображением. Легко видеть, что $\mathcal{A}'(x) = \Lambda$ и $\mathcal{A}''(x) = 0$ для любого $x \in X$.

Пример 3. *Оператор Немыцкого.* Пусть $[t_0, t_1]$ — конечный отрезок, G открытое подмножество $\mathbb{R} \times \mathbb{R}^n$ и $f\colon G \to \mathbb{R}^m$ — функция переменных $t \in \mathbb{R}$ и $x \in \mathbb{R}^n$. Пусть, далее, существует функция $x(\cdot) \in C([t_0, t_1], \mathbb{R}^n)$ такая, что $\Gamma(x(\cdot)) = \{\, (t, x(t)) \mid t \in [t_0, t_1] \,\} \subset G$. Легко проверить, что множество $U = \{\, x(\cdot) \in C([t_0, t_1], \mathbb{R}^n) \mid \Gamma(x(\cdot)) \subset G \,\}$ открыто в $C([t_0, t_1], \mathbb{R}^n)$. Пусть f непрерывна на G. Отображение $F\colon U \to C([t_0, t_1], \mathbb{R}^n)$, определенное по правилу

$$F(x(\cdot))(t) = f(t, x(t)), \quad \forall\, t \in [t_0, t_1],$$

называется *оператором Немыцкого*.

Предложение 1. *Пусть функция f и ее частная производная по x непрерывны на G. Тогда оператор Немыцкого F непрерывно дифференцируем на U и при этом, $F'(x(\cdot))[h(\cdot)](t) = f_x(t, x(t))h(t)$ для любых $x(\cdot) \in U$, $h(\cdot) \in C([t_0, t_1], \mathbb{R}^n)$ и $t \in [t_0, t_1]$.*

Доказательство. Пусть $\widehat{x}(\cdot) \in U$. Существует $\delta_0 > 0$ такое, что компакт $K = \{\, (t, x) \mid t \in [t_0, t_1],\ |x - \widehat{x}(t)| \le \delta_0 \,\}$ принадлежит G.

Пусть $\varepsilon > 0$. Функция f_x равномерно непрерывна на K и поэтому найдется $0 < \delta \le \delta_0$ такое, что если $|x_1 - x_2| < \delta$, то $\|f_x(t, x_1) - f_x(t, x_2)\| < \varepsilon$ для всех $(t, x_i) \in K$, $i = 1, 2$.

Для любого $t \in [t_0, t_1]$ отображение $g\colon B_{\mathbb{R}^n}(\widehat{x}(t), \delta) \to \mathbb{R}^m$, $g(x) = f(t, x) - f_x(t, \widehat{x}(t))x$, дифференцируемо на $B_{\mathbb{R}^n}(\widehat{x}(t), \delta)$ и его производная в точке x имеет вид: $g'(x) = f_x(t, x) - f_x(t, \widehat{x}(t))$.

Пусть $x_i(\cdot) \in B_{C([t_0, t_1], \mathbb{R}^n)}(\widehat{x}(\cdot), \delta)$, $i = 1, 2$. Тогда $x_i(t) \in B_{\mathbb{R}^n}(\widehat{x}(t), \delta)$, $i = 1, 2$, и мы имеем по теореме о среднем, примененной к отображению g (учитывая, что если $x \in [x_1(t), x_2(t)]$, то $x \in B_{\mathbb{R}^n}(\widehat{x}(t), \delta)$)

$$\begin{aligned}
&|f(t, x_1(t)) - f(t, x_2(t)) - f_x(t, \widehat{x}(t))(x_1(t) - x_2(t))| \le \\
&\le \sup_{x \in [x_1(t), x_2(t)]} \|f_x(t, x) - f_x(t, \widehat{x}(t))\| |x_1(t) - x_2(t)| \le \varepsilon |x_1(t) - x_2(t)|.
\end{aligned}$$

Поскольку это верно для любого $t \in [t_0, t_1]$, то отсюда следует, что отображение F строго дифференцируемо в $\widehat{x}(\cdot)$ и что $F'(\widehat{x}(\cdot)) = f_x(\cdot, \widehat{x}(\cdot))$. Так как $\widehat{x}(\cdot)$ — произвольная функция из U, то F дифференцируемо на U.

Непрерывная дифференцируемость отображения F на U проверяется непосредственно, используя снова равномерную непрерывность f_x на соответствующем компакте и определение нормы оператора. $\square$

Нам понадобится еще один оператор, который является небольшим обобщением оператора Немыцкого. Пусть $[t_0, t_1]$ — конечный отрезок, G — открытое подмножество $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r$ и $f\colon G \to \mathbb{R}^m$ — функция переменных $t \in \mathbb{R}$, $x \in \mathbb{R}^n$ и $u \in \mathbb{R}^r$, Пусть, далее, существует пара $(x(\cdot), u(\cdot)) \in C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$ такая, что $\Gamma(x(\cdot), u(\cdot)) = \{\, (t, x(t), u(t)) \mid t \in [t_0, t_1] \,\} \subset G$. Легко проверить, что множество $U = \{\, (x(\cdot), u(\cdot)) \in C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r) \mid \Gamma(x(\cdot), u(\cdot)) \subset G \,\}$ открыто в $C([t_0, t_1], \mathbb{R}^n)$. Пусть функция f непрерывна на G. Определим отображение $F\colon U \to C([t_0, t_1], \mathbb{R}^n)$ по правилу

$$F(x(\cdot), u(\cdot))(t) = f(t, x(t), u(t)), \quad \forall\, t \in [t_0, t_1],$$

которое назовем *обобщенным оператором Немыцкого.*

Следствие. *Пусть функция f и ее частные производные по x и u непрерывны на G. Тогда обобщенный оператор Немыцкого F непрерывно дифференцируем на U и при этом, $F'(x(\cdot), u(\cdot))[h(\cdot), \xi(\cdot)](t) = f_x(\cdot, x(\cdot), u(\cdot))h(t) + f_u(\cdot, x(\cdot), u(\cdot))\xi(t)$ для любых $(x(\cdot), u(\cdot)) \in U$, $(h(\cdot), \xi(\cdot)) \in C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$ и $t \in [t_0, t_1]$.*

Доказательство. Частная производная по $x(\cdot)$ отображения F, согласно предложению, непрерывна на U как отображение из $C([t_0, t_1], \mathbb{R}^n)$ и значит, тем более непрерывно как отображение из $C^1([t_0, t_1], \mathbb{R}^n)$. Частная производная по $u(\cdot)$ также непрерывна и поэтому по теореме о полном дифференциале отображение F непрерывно дифференцируемо на U и справедлива соответствующая формула для производной. $\square$

1.2.3. Оценка расстояния до поверхности уровня отображения. Теорема Люстерника. В этом пункте будет доказана теорема, являющаяся основным инструментом при получении необходимых условий экстремума.

Теорема (о поправке). *Пусть X и Y — банаховы пространства, U — окрестность точки $\widehat{x} \in X$, отображение $F\colon U \to Y$ строго дифференцируемо в $\widehat{x}$ и $\operatorname{Im} F'(\widehat{x}) = Y$. Тогда найдутся окрестность $U_0 \subset U$ точки $\widehat{x}$, отображение $g\colon U_0 \to X$ и константа $K > 0$ такие, что*

$$F(x + g(x)) = F(\widehat{x}) \tag{1}$$

и

$$\|g(x)\|_X \le K\|F(x) - F(\widehat{x})\|_Y \tag{2}$$

для всех $x \in U_0$.

Доказательство. Обозначим для краткости $\Lambda = F'(\widehat{x})$. Так как $\operatorname{Im} F'(\widehat{x}) = Y$, то по лемме о правом обратном (см. п. 1.4.1) существует отображение $R\colon Y \to X$ и константа $\gamma > 0$ такие, что $\Lambda(R(y)) = y$ и $\|R(y)\|_X \le \gamma\|y\|_Y$ для всех $y \in Y$.

Пусть $\varepsilon > 0$ таково, что $\theta = \varepsilon\gamma < 1$. Так как отображение F строго дифференцируемо в $\widehat{x}$, то найдется такое $\delta > 0$, что $B_X(\widehat{x}, \delta) \subset U$ и для любых $x, x' \in B_X(\widehat{x}, \delta)$ справедливо неравенство

$$\|F(x) - F(x') - \Lambda(x - x')\|_Y \le \varepsilon\|x - x'\|_X = \frac{\theta}{\gamma}\|x - x'\|_X. \tag{i}$$

Выберем окрестность U_0 так, что $U_0 \subset B_X(\widehat{x}, \delta/2)$ и если $x \in U_0$, то $\|F(x) - F(\widehat{x})\|_Y < \delta(1 - \theta)/2\gamma$.

Пусть $x \in U_0$. Рассмотрим последовательность (модифицированный метод Ньютона)

$$x_n = x_{n-1} - R(F(x_{n-1}) - F(\widehat{x})), \quad n \in \mathbb{N}, \quad x_0 = x. \tag{ii}$$

Докажем, что эта последовательность принадлежит $B_X(\widehat{x}, \delta)$ и фундаментальна. Первое доказываем по индукции. Ясно, что $x_0 \in B_X(\widehat{x}, \delta)$. Пусть $x_k \in B_X(\widehat{x}, \delta)$, $1 \le k \le n$. Применяя к обеим частям (ii) оператор Λ, получим

$$\Lambda(x_n - x_{n-1}) = -F(x_{n-1}) + F(\widehat{x}). \tag{iii}$$

Используя последовательно (ii), оценку для правого обратного, (iii), (i) и затем итерируя процедуру, будем иметь

$$\begin{aligned} &\|x_{n+1} - x_n\|_X \le \gamma\|F(x_n) - F(\widehat{x})\|_Y = \gamma\|F(x_n) - F(x_{n-1}) - \\ &\quad - \Lambda(x_n - x_{n-1})\|_Y \le \theta\|x_n - x_{n-1}\|_X \le \ldots \le \theta^n\|x_1 - x\|_X. \end{aligned} \tag{iv}$$

Далее по неравенству треугольника, (iv), (ii), условию 2) теоремы и согласно определению окрестности U_0, получаем, что

$$\begin{aligned} &\|x_{n+1} - \widehat{x}\|_X \le \|x_{n+1} - x\|_X + \|x - \widehat{x}\|_X \le \|x_{n+1} - x_n\|_X + \ldots \\ &\quad \ldots + \|x_1 - x\|_X + \|x - \widehat{x}\|_X \le (\theta^n + \theta^{n-1} + \ldots + 1)\|x_1 - x\|_X + \\ &+ \|x - \widehat{x}\|_X \le \frac{\gamma}{1 - \theta}\|F(x) - F(\widehat{x})\|_Y + \|x - \widehat{x}\|_X < \frac{\delta}{2} + \frac{\delta}{2} = \delta, \end{aligned} \tag{v}$$

т. е. $x_{n+1} \in B_X(\widehat{x}, \delta)$ и значит, вся последовательность $\{x_n\}$ принадлежит $B_X(\widehat{x}, \delta)$.

Последовательность $\{x_n\}$ фундаментальна. Действительно, используя (iv) и рассуждая как в предыдущем неравенстве, будем иметь для всех $n, m \in \mathbb{N}$

$$\|x_{n+m} - x_n\|_X \le \|x_{n+m} - x_{n+m-1}\|_X + \ldots + \|x_{n+1} - x_n\|_X \le$$
$$\le (\theta^{n+m-1} + \ldots + \theta^n)\|x_1 - x\|_X \le \frac{\theta^n}{1-\theta}\|x_1 - x\|_X \le \delta\theta^n.$$

Положим $g(x) = \lim_{n\to\infty} x_n - x$. Из (v) следует, что $x + g(x) \in B_X(\widehat{x}, \delta) \subset V$. Переходя к пределу в (iii) при $n \to \infty$, получаем, что $F(x + g(x)) = F(\widehat{x})$.

Из (v) при $\widehat{x} = x$ вытекает, что $\|x_n - x\|_X \le (\gamma/(1-\theta))\|F(x) - F(\widehat{x})\|_Y$. Переходя здесь к пределу при $n \to \infty$, приходим к неравенству $\|g(x)\|_X \le K\|F(x) - F(\widehat{x})\|_Y$, где $K = \gamma/(1-\theta)$. □

Определение. *Пусть M — непустое подмножество нормированного пространства X. Элемент $h \in X$ называется касательным вектором к M в точке $\widehat{x} \in M$, если существуют $\varepsilon > 0$ и отображение $r\colon (-\varepsilon, \varepsilon) \to X$ такие, что $\widehat{x} + th + r(t) \in M$ для всех $t \in (-\varepsilon, \varepsilon)$ и $\|r(t)\|_X/t \to 0$, при $t \to 0$. Множество всех касательных векторов к M в точке $\widehat{x} \in M$ обозначим $T_{\widehat{x}}M$.*

Следствие (Теорема Люстерника)**.** *Пусть выполнены условия теоремы о поправке и $M = \{\, x \in U \mid F(x) = F(\widehat{x}) \,\}$. Тогда $T_{\widehat{x}}M = \operatorname{Ker} F'(\widehat{x})$.*

Доказательство. Пусть $h \in T_{\widehat{x}}M$ и r из определения касательного вектора. Тогда вследствие дифференцируемости F в точке $\widehat{x}$ имеем $0 = F(\widehat{x} + th + r(t)) - F(\widehat{x}) = tF'(\widehat{x})h + o(t)$, откуда (деля на t и переходя к пределу при $t \to 0$) следует, что $h \in \operatorname{Ker} F'(\widehat{x})$. Обратно, пусть $h \in \operatorname{Ker} F'(\widehat{x})$ и U_0, g и K из теоремы о поправке. Существует $\varepsilon > 0$ такое, что $\widehat{x} + th \in V_0$, если $t \in (-\varepsilon, \varepsilon)$. Положим $r(t) = g(\widehat{x} + th)$. Тогда $F(\widehat{x} + th + r(t)) = F(\widehat{x})$ и $\|r(t)\|_X \le K\||F(\widehat{x} + th) - F(\widehat{x})\|_Y = K\|F(\widehat{x}) + tF'(\widehat{x})h + o(t) - F(\widehat{x})\|_Y = K\|o(t)\|_Y$, т. е. h — касательный вектор. □

Глава 2. Условия экстремума для функций на банаховых пространствах

Пусть X — банахово пространство, $f\colon X \to \mathbb{R}$ и C — непустое подмножество X. В этой главе изучается следующая задача на минимум

$$f(x) \to \min, \quad x \in C \tag{P_0}$$

для различных множеств C, задающих вид ограничений.

Задача на максимум сводится к задаче на минимум, если заменить функцию f на $-f$.

§ 2.1. Гладкие задачи без ограничений

2.1.1. Постановка задачи. Если $C = X$ в (P_0), то мы имеем задачу

$$f(x) \to \min, \quad x \in X \qquad (P_1),$$

которая называется *задачей без ограничений.* Если функция f обладает некоторой гладкостью, то говорят о *гладкой задаче без ограничений.*

Точка $\widehat{x}$ называется *локальным минимумом* в задаче (P_1), если существует такая окрестность U этой точки, что $f(x) \geq f(\widehat{x})$ для всех $x \in U$.

2.1.2. Теорема Ферма. Необходимые и достаточные условия второго порядка.

Теорема (Теорема Ферма — необходимые условия минимума I-го порядка в задаче без ограничений)**.** *Если $\widehat{x}$ — локальный минимум в задаче (P_1) и функция f дифференцируема в $\widehat{x}$, то*

$$f'(\widehat{x}) = 0. \tag{1}$$

Доказательство. Пусть $h \in X$. Так как $\widehat{x}$ — локальный минимум и функция f дифференцируема в $\widehat{x}$, то для достаточно малых $t > 0$ имеем

$$0 \leq f(\widehat{x} + th) - f(\widehat{x}) = t\langle f'(\widehat{x}), h\rangle + o(t).$$

Деля на t и переходя к пределу при $t \to 0$, получаем, что $\langle f'(\widehat{x}), h\rangle \geq 0$, т. е. линейный функционал $f'(\widehat{x})$ неотрицателен на всем пространстве и значит, он нулевой ($\langle f'(\widehat{x}), h\rangle = 0$ для любого $h \in X$). □

В задаче на максимум, очевидно, те же необходимые условия.

Теорема (Необходимые условия минимума II-го порядка в задаче без ограничений)**.** *Если $\widehat{x}$ — локальный минимум в задаче (P_1) и функция f дважды дифференцируема в $\widehat{x}$, то*

$$f'(\widehat{x}) = 0, \qquad f''(\widehat{x})[h, h] \geq 0, \ \ \forall\ h \in X. \tag{2}$$

Доказательство. Первое условие в (2) следует из предыдущей теоремы. Пусть $h \in X$. Так как $\widehat{x}$ — локальный минимум, функция f дважды дифференцируема в $\widehat{x}$ и $f'(\widehat{x}) = 0$, то для достаточно малых $t > 0$ имеем по формуле Тейлора

$$0 \leq f(\widehat{x} + th) - f(\widehat{x}) = t^2 f''(\widehat{x})[h, h] + o(t^2).$$

Деля на t и переходя к пределу при $t \to 0$, получаем второе соотношение в (2). □

Теорема (Достаточные условия минимума II-го порядка в задаче без ограничений)**.** *Если в задаче (P_1) функция f дважды дифференцируема в $\widehat{x}$, $f'(\widehat{x}) = 0$ и существует $\alpha > 0$ такое, что*

$$f''(\widehat{x})[h, h] \geq \alpha \|h\|_X^2, \ \ \forall\ h \in X, \tag{3}$$

то $\widehat{x}$ — локальный минимум в задаче (P_1).

Доказательство. В силу условий теоремы для достаточно малых по норме h по формуле Тейлора

$$f(\widehat{x}+h)-f(\widehat{x})=f''(\widehat{x})[h,h]+o(\|h\|_X^2)\geq\frac{\alpha}{2}\|h\|_X^2+o(\|h\|_X^2).$$

Можно считать, что $|o(\|h\|_X^2)|<(\alpha/2)\|h\|_X^2$, и тогда $f(\widehat{x}+h)-f(\widehat{x})>0$ для таких h, т. е. $\widehat{x}$ — локальный минимум. □

§ 2.2. Гладкие задачи с ограничениями типа равенств

2.2.1. Постановка задачи. Пусть X и Y — банаховы пространства, $f\colon X\to\mathbb{R}$ и $F\colon X\to Y$. Задачу

$$f(x)\to\min,\quad F(x)=0 \qquad (P_2)$$

называют *задачей с ограничениями типа равенств*. Если функция f и отображение F обладают некоторой гладкостью, то говорят о *гладкой задаче с ограничениями типа равенств*.

Сопоставим задаче (P_2) *функцию Лагранжа*

$$\mathcal{L}(x,\lambda_0,y^*)=\lambda_0 f(x)+\langle y^*,F(x)\rangle,$$

где $\lambda_0\in\mathbb{R}$ и $y^*\in Y^*$ называются *множителями Лагранжа.*

2.2.2. Правило множителей Лагранжа.

Теорема (Правило множителей Лагранжа — необходимые условия минимума I-го порядка в задаче с ограничениями типа равенств). *Если $\widehat{x}$ — локальный минимум в задаче (P_2) и функция f дифференцируема в $\widehat{x}$, отображение F строго дифференцируемо в $\widehat{x}$ и $\operatorname{Im}F'(\widehat{x})$ — замкнутое подпространство в Y, то найдутся, не равные одновременно нулю, множители Лагранжа λ_0 и y^* такие, что*

$$\mathcal{L}_x(\widehat{x},\lambda_0,y^*)=0 \ \Leftrightarrow\ \lambda_0 f'(\widehat{x})+(F'(\widehat{x}))^*y^*=0. \qquad (1)$$

Если $\operatorname{Im}F'(\widehat{x})=Y$, *то* $\lambda_0\neq 0$.

Доказательство. Рассмотрим отдельно два случая.

A) Регулярный случай: $\operatorname{Im}F'(\widehat{x})=Y$. Пусть $h\in\operatorname{Ker}F'(\widehat{x})$. Отображение F удовлетворяет условиям теоремы Люстерника (см. ...) и поэтому $h\in T_{\widehat{x}}M$, т. е. существуют $\varepsilon>0$ и отображение $r\colon(-\varepsilon,\varepsilon)\to X$ такие, что $F(\widehat{x}+th+r(t))=0$ для $t\in(-\varepsilon,\varepsilon)$ и $r(t)=o(t)$ при $t\to 0$. Таким образом, элементы $\widehat{x}+th+r(t)$, $t\in(-\varepsilon,\varepsilon)$, допустимы в (P_2) и так как $\widehat{x}$ — локальный минимум в этой задаче, то $0\leq f(\widehat{x}+th+r(t))-f(\widehat{x})=t\langle f'(\widehat{x}),h\rangle+o(t)$ для достаточно малых t. Деля последнее соотношение на $t>0$ и устремляя t к нулю, получаем, что $\langle f'(\widehat{x}),h\rangle\geq 0$. Но h — произвольный элемент из $\operatorname{Ker}F'(\widehat{x})$ и поэтому $\langle f'(\widehat{x}),h\rangle=0$ для любого $h\in\operatorname{Ker}F'(\widehat{x})$, т. е. $f'(\widehat{x})\in(\operatorname{Ker}F'(\widehat{x}))^{\perp}$. Согласно лемме об аннуляторе ядра эпиморфизма (см. 1.1.4) $f'(\widehat{x})\in\operatorname{Im}(F'(\widehat{x}))^*$ и, следовательно, существует функционал $y^*\in(F'(\widehat{x}))^*$ такой, что $f'(\widehat{x})=-(F'(\widehat{x}))^*y^*$, или $f'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0$. В регулярном случае соотношение (1) с $\lambda_0 = 1$ доказано.

B) Нерегулярный (вырожденный) случай: $\operatorname{Im} F'(\widehat{x}) \neq Y$. Так как по условию подпространство $\operatorname{Im} F'(\widehat{x})$ замкнуто, то по лемме о нетривиальном аннулятора (см. п. 1.1.3) существует ненулевой функционал $y^* \in (F'(\widehat{x}))^*$ такой, что $\langle y^*, F'(\widehat{x})x \rangle = 0$ для любого $x \in X$, т. е. $(F'(\widehat{x}))^* y^* = 0$. Это доказывает утверждение теоремы в нерегулярном случае с $\lambda_0 = 0$. $\square$

Рассмотрим частный случай задачи (P_2), когда $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, а отображение F задается функциями $f_i \colon \mathbb{R}^n \to \mathbb{R}^m$, $i = 1, \ldots, m$, т. е. $F(x) = (f_1(x), \ldots, f_m(x))^T$, $x \in \mathbb{R}^n$. Таким образом, рассматривается задача

$$f_0(x) \to \min, \quad f_i(x) = 0, \quad i = 1, \ldots, m. \qquad (P_2')$$

Поскольку линейные функционалы на $\mathbb{R}^m$ суть вектор-строки $(\lambda_1, \ldots, \lambda_m)$ (см. п. 1.1.2), то функция Лагранжа задачи (P_2') имеет вид $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^m \lambda_i f_i(x)$, где $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m)$.

Следующее утверждение — классическое правило множителей Лагранжа для гладких конечномерных задач.

Следствие (Правило множителей Лагранжа)**.** *Если $\widehat{x}$ — локальный минимум в задаче (P_2') и функция f_0 дифференцируема в $\widehat{x}$, а функции f_i, $i = 1, \ldots, m$, строго дифференцируемы в $\widehat{x}$, то найдутся, не равные одновременно нулю, множители Лагранжа $\lambda_0, \lambda_1, \ldots, \lambda_m$ такие, что*

$$\mathcal{L}_x(\widehat{x}, \overline{\lambda}) = 0 \;\Leftrightarrow\; \sum_{i=0}^m \lambda_i f_i'(\widehat{x}) = 0.$$

Если векторы $f_1'(\widehat{x}), \ldots, f_m'(\widehat{x})$ линейно независимы, то $\lambda_0 \neq 0$.

Доказательство сразу следует из предыдущей теоремы, если учесть, что подпространство $\operatorname{Im} F'(\widehat{x})$ конечномерно и поэтому замкнуто, а линейная независимость векторов $f_1'(\widehat{x}), \ldots, f_m'(\widehat{x})$ означает, что $\operatorname{Im} F'(\widehat{x}) = Y$.

2.2.3. Условия экстремума второго порядка.

Теорема (Необходимые условия минимума II-го порядка в задаче с ограничениями типа равенств)**.** *Если $\widehat{x}$ — локальный минимум в задаче (P_2), а функция f и отображение F дважды дифференцируемы в $\widehat{x}$ и $\operatorname{Im} F'(\widehat{x}) = Y$, то найдется множитель Лагранжа $y^* \in Y^*$ такой, что*

$$\mathcal{L}_x(\widehat{x}, 1, y^*) = 0 \;\Leftrightarrow\; f'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0 \qquad (2)$$

и

$$\mathcal{L}_{xx}(\widehat{x},1,y^*)[h,h] \geq 0, \ \ \forall\ h \in \operatorname{Ker} F'(\widehat{x}) \ \Leftrightarrow$$
$$f''(\widehat{x})[h,h] + \langle y^*, F''(\widehat{x})[h,h]\rangle \geq 0, \ \ \forall\ h \in \operatorname{Ker} F'(\widehat{x}). \quad (3)$$

Доказательство. Из дважды дифференцируемости f и F в точке $\widehat{x}$ следует непрерывность производных и $\widehat{x}$ и, следовательно, строгая дифференцируемость этих отображений в данной точке (см. следствие п. 1.2.1). В частности функция f дифференцируема в $\widehat{x}$ и поэтому соотношение (2) (с $\lambda_0 = 1$) сразу следует из предыдущей теоремы.

Докажем (3). Пусть $h \in \operatorname{Ker} F'(\widehat{x})$. Рассуждая дословно также как в предыдущей теореме, находим допустимые в (P_2) точки $\widehat{x} + th + r(t)$, $t \in (-\varepsilon, \varepsilon)$, где $r(t) = o(t)$ при $t \to 0$, и так как $\widehat{x}$ — локальный минимум, то $f(\widehat{x} + th + r(t)) \geq f(\widehat{x})$ для достаточно малых t. Теперь по формуле Тейлора имеем (учитывая (2))

$$0 \leq f(\widehat{x} + th + r(t)) - f(\widehat{x}) = \mathcal{L}(\widehat{x} + th + r(t), 1, y^*) - \mathcal{L}(\widehat{x}, 1, y^*) =$$
$$= \frac{1}{2}\mathcal{L}_{xx}(\widehat{x},1,y^*)[th+r(t), th+r(t)] + o(t^2) = \frac{t^2}{2}\mathcal{L}_{xx}(\widehat{x},1,y^*)[h,h] + o(t^2),$$

откуда следует (3). □

Теорема (Достаточные условия минимума II-го порядка в задаче с ограничениями типа равенств)**.** *Пусть в задаче* (P_2) *функция* f *и отображение* F *дважды дифференцируемы в допустимой точке* $\widehat{x}$ *и* $\operatorname{Im} F'(\widehat{x})$ *— замкнутое подпространство в* Y*. Тогда, если найдутся множитель Лагранжа* $y^* \in Y^*$ *и число* $\alpha > 0$ *такие, что*

$$\mathcal{L}_x(\widehat{x}, 1, y^*) = 0 \quad (4)$$

и

$$\mathcal{L}_{xx}(\widehat{x},1,y^*)[h,h] \geq \alpha\|h\|_X^2, \ \ \forall\ h \in \operatorname{Ker} F'(\widehat{x}), \quad (5)$$

то $\widehat{x}$ *— локальный минимум в задаче* (P_2)*.*

Доказательство. Рассмотрим отображение $G\colon X \to \operatorname{Im} F'(\widehat{x})$, определенное по формуле $G(x) = F'(\widehat{x})x$. Для него, очевидно, выполнены условия теоремы о поправке, согласно которой существуют окрестность V_0 точки ноль, отображение $g\colon V_0 \to X$ и константа $K > 0$ такие, что

$$F'(\widehat{x})(x + g(x)) = 0 \quad (i)$$

и

$$\|g(x)\|_X \leq K\|F'(x)x\|_Y \quad (ii)$$

для всех $x \in V_0$.

Можно считать, что в любой окрестности $\widehat{x}$ есть допустимые в (P_2) точки (иначе $\widehat{x}$ была бы изолированной точкой и автоматически локальным минимумом). Пусть $x \in V_0$ и $\widehat{x} + x$ — допустимый элемент в задаче (P_2). Тогда по формуле Тейлора

$$0 = F(\widehat{x}+x) = F'(\widehat{x})x + \frac{1}{2}F''(\widehat{x})[x,x] + o(\|x\|_X^2).$$

отсюда следует, что для достаточно малых x справедливо неравенство $\|F'(\widehat{x})x\|_Y \leq ((1/2)\|F''(\widehat{x})\|+1)\|x\|_X^2$, а тогда из (ii) получаем, что

$$\|g(x)\|_X \leq \gamma\|x\|_X^2, \qquad (iii)$$

где $\gamma = K((1/2)\|F''(\widehat{x})\|+1)\|x\|_X^2$.

Обозначая, для краткости, $L(x) = \mathcal{L}(x,1,y^*)$, снова по формуле Тейлора получаем (учитывая (1) и то, что $\widehat{x}+x$ — допустимая точка)

$$f(\widehat{x}+x) = f(\widehat{x}) + \frac{1}{2}L_{xx}(\widehat{x})[x,x] + o(\|x\|_X^2).$$

Отсюда, полагая $B = \|L_{xx}(\widehat{x})\|$, используя (2) (учитывая, что $x+g(x) \in \operatorname{Ker} F'(\widehat{x})$ согласно (i)) и полученные выше оценки, будем иметь

$$f(\widehat{x}+x)-f(\widehat{x}) = \frac{1}{2}L_{xx}(\widehat{x})[x+g(x)-g(x),x+g(x)-g(x)]+o(\|x\|_X^2) =$$
$$= \frac{1}{2}(L_{xx}(\widehat{x})[x+g(x),x+g(x)] - 2L_{xx}(\widehat{x})[x+g(x),g(x)]+$$
$$+ L_{xx}(\widehat{x})[g(x),g(x)]) + o(\|x\|_X^2) \geq \frac{1}{2}(\alpha\|x+g(x)\|_X^2-$$
$$- 2B\|x+g(x)\|_X\|g(x)\|_X - B\|g(x)\|_X^2) + o(\|x\|_X^2).$$

Нетрудно проверить, что в силу оценки (iii) выражение справа положительно для достаточно малых x и поэтому $\widehat{x}$ — локальный минимум. $\square$

Заметим, что доказано чуть больше: существует такое $\kappa > 0$ и окрестность нуля V в X, что $f(\widehat{x}+x) - f(\widehat{x}) \geq \kappa\|x\|_X^2$ для всех таких $x \in V$, для которых точка $\widehat{x}+x$ допустима в задаче (P_2).

Отметим также, что условие (5) в теореме и непрерывность билинейного отображения $(h_1,h_2) \mapsto \mathcal{L}_{xx}(\widehat{x},1,y^*)[h_1,h_2]$ означают, что это билинейное отображение определяет скалярное произведение на $\operatorname{Ker} F'(\widehat{x})$ и соответствующая норма эквивалентна исходной. Другими словами, данная теорема применима, по сути дела, лишь для евклидовых пространств. Мы сейчас приведем обобщение этого результата на более широкий класс пространств. Сначала некоторые определения.

Пусть X и X_1 — нормированные пространства. Говорят, что X *непрерывно вложено в* X_1, если $X \subset X_1$ и существует такая константа $c > 0$, что $\|x\|_{X_1} \leq c\|x\|_X$ для любого $x \in X$.

Определение. *Пусть X, X_1 и Y — нормированные пространства, X непрерывно вложено в X_1 и U — окрестность точки*

$\widehat{x} \in X$. Будем говорить, что для отображения $\Phi\colon U \to Y$ справедливо разложение второго порядка в точке $\widehat{x}$ относительно пары (X, X_1), если Φ дважды дифференцируемо в $\widehat{x}$, вторая производная $\Phi''(\widehat{x})$ непрерывна, как билинейная форма, на $X_1 \times X_1$ и для всех $x \in X$, для которых $\widehat{x} + x \in U$ справедливо представление

$$\Phi(\widehat{x} + x) = \Phi(\widehat{x}) + \Phi'(\widehat{x})x + \frac{1}{2}\Phi''[x, x] + r(x),$$

где $\|r(x)\|_Y / \|x\|^2_{X_1} \to 0$ при $x \to 0$ в X.

Теорема (Обобщенные достаточные условия минимума II-го порядка в задаче с ограничениями типа равенств)**.** *Пусть в задаче (P_2) пространство X непрерывно вложено в банахово пространство X_1, для функционала f и отображения F справедливы разложения второго порядка в точке $\widehat{x}$ относительно пары (X, X_1) и $\operatorname{Im} F'(\widehat{x})$ — замкнутое подпространство. Тогда, если найдутся множитель Лагранжа $y^* \in Y^*$ и число $\alpha > 0$ такие, что*

$$\mathcal{L}_x(\widehat{x}, 1, y^*) = 0$$

и

$$\mathcal{L}_{xx}(\widehat{x}, 1, y^*)[h, h] \geq \alpha \|h\|^2_{X_1}, \ \ \forall\, h \in \operatorname{Ker} F'(\widehat{x}),$$

то $\widehat{x}$ — локальный минимум в задаче (P_2).

Для доказательства этого утверждения надо практически дословно повторить рассуждения в предыдущей теореме, заменяя разложение по формуле Тейлора разложением второго порядка.

§ 2.3. Гладкие задачи с ограничениями типа равенств и неравенств

2.3.1. Постановка задачи. Пусть X и Y — банаховы пространства, $f_i\colon X \to \mathbb{R}$, $i = 0, 1, \dots, m$, и $F\colon X \to Y$. Задачу

$$f_0(x) \to \min, \quad f_i(x) \leq 0, \quad 1 \leq i \leq m, \quad F(x) = 0 \qquad (P_3)$$

называют *задачей с ограничениями типа равенств и неравенств*. Если функция f_i, $i = 0, 1, \dots, m$, и отображение F обладают некоторой гладкостью, то говорят о *гладкой задаче с ограничениями типа равенств и неравенств*.

Сопоставим задаче (P_3) *функцию Лагранжа*

$$\mathcal{L}(x, \overline{\lambda}, y^*) = \sum_{i=0}^{m} \lambda_i f_i(x) + \langle y^*, F(x) \rangle,$$

где $\overline{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и $y^* \in Y^*$. Числа λ_i, $i = 0, 1, \dots, m$, и функционал y^* называются *множителями Лагранжа*.

2.3.2. Правило множителей Лагранжа.

Теорема (Правило множителей Лагранжа — необходимые условия минимума I-го порядка в задаче с ограничениями типа равенств и неравенств)**.** *Если $\widehat{x}$ — локальный минимум в задаче (P_3), а функции f_i, $i = 0, 1, \ldots, m$, и отображение F строго дифференцируемы в $\widehat{x}$ и $\operatorname{Im} F'(\widehat{x})$ — замкнутое подпространство в Y, то найдутся такие множители Лагранжа $\lambda_0, \lambda_1, \ldots, \lambda_m$ и y^*, не равные нулю одновременно, что выполняются*

- (a) $\mathcal{L}_x(\widehat{x}, \overline{\lambda}) = 0 \Leftrightarrow \sum_{i=0}^{m} \lambda_i f_i'(\widehat{x}) + (F'(\widehat{x}))^* y^* = 0$ *(условие стационарности);*
- (b) $\lambda_i \geq 0,\ i = 0, 1, \ldots, m$ *(условие неотрицательности);*
- (c) $\lambda_i f_i(\widehat{x}) = 0,\ i = 1, \ldots, m$ *(условие дополняющей нежесткости).*

Если $\operatorname{Im} F'(\widehat{x}) = Y$ и существует вектор $h \in \operatorname{Ker} F'(\widehat{x})$ такой, что $\langle f_i'(\widehat{x}), h\rangle < 0$ для тех $1 \leq i \leq m$, для которых $f_i(\widehat{x}) = 0$, то $\lambda_0 \neq 0$.

Доказательство. Заметим сначала, что утверждение (c) можно считать выполненным всегда. В самом деле, если $f_i(\widehat{x}) < 0$ для некоторого $1 \leq i \leq m$, то это ограничение не существенно для локального минимума, поскольку будет выполняться в достаточно малой окрестности точки $\widehat{x}$. Поэтому такое ограничение можно отбросить и положить для него $\lambda_i = 0$.

Как и в доказательстве правила множителей Лагранжа для задачи с ограничениями типа равенств рассмотрим отдельно два случая.

A) Невырожденный случай: $\operatorname{Im} F'(\widehat{x}) = Y$. Рассмотрим множество

$$C = \{\, (\mu_0, \mu_1, \ldots, \mu_m, y) \in \mathbb{R}^{m+1} \times Y \mid \exists\, x \in X : \mu_i > \langle f_i'(\widehat{x}), x\rangle, \\ i = 0, 1, \ldots, m,\ y = F'(\widehat{x})x \,\}.$$

Для доказательства теоремы достаточно проверить, что C — выпуклое множество, $0 \notin C$ и $\operatorname{int} C \neq \emptyset$ и . Действительно, в этом случае согласно первой теоремы отделимости (см. п. 1.1.3), примененной ко множествам $\{0\}$ и C, найдется ненулевой функционал $(\lambda_0, \lambda_1, \ldots, \lambda_m, y^*) \in (\mathbb{R}^{m+1})^* \times Y^*$ такой, что

$$\sum_{i=0}^{m} \lambda_i \mu_i + \langle y^*, y\rangle \geq 0, \quad \forall\, (\mu_0, \mu_1, \ldots, \mu_m, y) \in C. \qquad (i)$$

Наборы $(\mu_0, \mu_1, \ldots, \mu_m, 0)$, где $\mu_i > 0$, $i = 0, 1 \ldots, m$, принадлежат C (надо взять $x = 0$). Подставляя их в (i), получаем, что $\lambda_i \geq 0$, $i = 0, 1 \ldots, m$, и утверждение (b) теоремы доказано.

Далее, для любого $x \in X$ и любого $\varepsilon > 0$ наборы $(\langle f_0'(\widehat{x}), x\rangle + \varepsilon, \ldots, \langle f_m'(\widehat{x}), x\rangle + \varepsilon, F'(\widehat{x})x)$, очевидно, также принадлежат C. Подставляя их в (i), получаем, что

$$\sum_{i=0}^{m} \lambda_i \langle f_i'(\widehat{x}), x\rangle + \langle y^*, F'(\widehat{x})x\rangle \geq -\varepsilon \sum_{i=0}^{m} \lambda_i.$$

В силу произвольности ε левая часть этого неравенства (которая есть линейный функционал) неотрицательна на X и значит, она равна нулю, а это равносильно утверждению (a) теоремы.

Итак, осталось проверить, что C — выпуклое множество, $0 \notin C$ и $\operatorname{int} C \neq \emptyset$. Первое проверяется без труда. Второе докажем от противного. Пусть $0 \in C$. Тогда существует такое $x_0 \in X$, что $\langle f_i'(\widehat{x}), x_0\rangle < 0$, $i = 0, 1, \ldots, m$, и $F'(\widehat{x})x_0 = 0$. По теореме Люстерника (см. п. 1.3.2) $x_0 \in T_{\widehat{x}}M$, где $M = \{\, x \in X \mid F(x) = F(\widehat{x}) = 0 \,\}$, т. е. существуют $\varepsilon > 0$ и отображение $r\colon (-\varepsilon, \varepsilon) \to X$ такие, что $F(\widehat{x} + tx_0 + r(t)) = 0$ для всех $t \in (-\varepsilon, \varepsilon)$ и $\|r(t)\|_X = o(t)$ при $t \to 0$. Далее, в силу дифференцируемости функций f_i, $i = 0, 1, \ldots, m$, в точке $\widehat{x}$ имеем $f_i(\widehat{x} + tx_0 + r(t)) = f_i(\widehat{x}) + \langle f_i'(\widehat{x}), x_0\rangle + o(t) < f_i(\widehat{x})$ для достаточно малых $t > 0$. Это значит, что для таких t точки $\widehat{x} + tx_0 + r(t)$ допустимы в задаче (P_3), а значение функционала f_0 на них меньше, чем $f_0(\widehat{x})$, в противоречие с тем, что $\widehat{x}$ — локальный минимум. Итак, $0 \notin C$.

Покажем теперь, что $\operatorname{int} C \neq \emptyset$. Рассмотрим множество $C_0 = \{\, (\mu_0, \mu_1, \ldots, \mu_m, y) \in \mathbb{R}^{m+1} \times Y \mid \mu_i > d,\ i = 0, 1, \ldots, m,\ y \in F'(\widehat{x})(B_X(0,1)) \,\}$, где $d = \max_{0 \leq i \leq m} \|f_i'(\widehat{x})\|_{X^*}$. Это множество открыто, потому что открыто множество $F'(\widehat{x})(B_X(0,1))$ в силу теоремы Банаха об открытом отображении (см. п. 1.1.4), и $C_0 \subset C$. Действительно, пусть $(\mu_0, \mu_1, \ldots, \mu_m, y) \in C_0$ и $x \in B_X(0,1)$ такое, что $y = F'(\widehat{x})x$. Тогда $\mu_i > d \geq \langle f'(\widehat{x}), x\rangle$, $i = 0, \ldots, m$, и значит, $(\mu_0, \mu_1, \ldots, \mu_m, y) \in C$, т. е. $\operatorname{int} C \neq \emptyset$.

B) Вырожденный случай: $\operatorname{Im} F'(\widehat{x}) \neq Y$. Здесь надо, фактически, дословно повторить доказательство правила множителей Лагранжа для вырожденного случая в гладкой задаче с равенствами.

Докажем последнее утверждение теоремы. Пусть выполнены его предположения и $\lambda_0 = 0$. Тогда из (a) последует, что $\lambda_i = 0$, $i = 1, \ldots, m$, и значит, $(F'(\widehat{x}))^* y^* =$, т. е. $y^* = 0$ (так как из сюръективности $F'(\widehat{x})$ следует инъективность сопряженного оператора, см. п. 1.1.1). Это противоречит тому, что не все множители Лагранжа равны нулю. $\square$

§ 2.4. Выпуклые задачи

Пусть X — вещественное векторное пространство. Напомним, что непустое множество $A \subset X$ называется *выпуклым*, если с любыми двумя своими точками оно содержит и отрезок $[x, y] = \{ z \in X \mid z = (1-\alpha)x + \alpha y,\ 0 \le \alpha \le 1 \}$, соединяющий эти точки.

Пусть $f\colon X \to \mathbb{R}$. Множество

$$\operatorname{epi} f = \{ (x, \alpha) \in X \times \mathbb{R} \mid \alpha \ge f(x) \}$$

называется *надграфиком* f.

Функция $f\colon X \to \mathbb{R}$ называется *выпуклой*, если ее надграфик — выпуклое множество.

Нетрудно проверить, что функция f выпукла тогда и только тогда, когда для любых $x_1, x_1 \in X$ и любого $0 \le \alpha \le 1$ выполняется неравенство

$$f((1-\alpha)x_1 + \alpha x_2) \le (1-\alpha)f(x_1) + \alpha f(x_2),$$

которое называется *неравенством Иенссена.*

2.4.1. Постановка задачи. Пусть X — вещественное векторное пространство, $f_i\colon X \to \mathbb{R}$, $i = 0, 1, \ldots, m$, — выпуклые функции и A — непустое выпуклое подмножество X. Задачу

$$f_0(x) \to \min, \quad f_i(x) \le 0, \quad i = 1, \ldots, m, \quad x \in A \qquad (P_4)$$

называют *выпуклой задачей* или *задачей выпуклого программирования.*

Отметим, что функции определены на векторном пространстве и поэтому можно говорить только о глобальном минимуме. Но даже если бы X было нормированным пространством, то легко показать, что для выпуклой задачи любой локальный минимум является глобальным.

Свяжем с задачей (P_4) следующую функцию Лагранжа: $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^{m} \lambda_i f_i(x)$, где $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ — набор множителей Лагранжа.

2.4.2. Теорема Каруша–Куна–Таккера.

Теорема (Каруша–Куна–Таккера)**.** *Если $\widehat{x}$ — минимум в задаче (P_4), то найдется такой ненулевой набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, что выполнены*

- (a) $\min_{x\in A} \mathcal{L}(x, \overline{\lambda}) = \mathcal{L}(\widehat{x}, \overline{\lambda})$ *(условие минимума);*
- (b) $\lambda_i \ge 0,\ \ i = 0, 1, \ldots, m$ *(условие неотрицательности);*
- (c) $\lambda_i f_i(\widehat{x}) = 0,\ \ i = 1, \ldots, m$ *(условие дополняющей нежесткости).*

Если существует допустимая в (P_4) точка $\widehat{x}$ и набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, удовлетворяющие условиям (a), (b) и (c) и при этом $\lambda_0 > 0$, то $\widehat{x}$ — решение задачи (P_4).

Если найдется точка $\overline{x} \in A$ такая, что $f_i(\overline{x}) < 0$, $1 \le i \le m$, то $\lambda_0 \ne 0$ (условие Слейтера).

Доказательство. Пусть $\widehat{x}$ — решение задачи (P_5). Рассмотрим множество $M = \{\, \mu = (\mu_0, \mu_1, \ldots, \mu_m)^T \in \mathbb{R}^{m+1} \mid \exists\, x \in A : f_0(x) - f_0(\widehat{x}) < \mu_0,\ f_i(x) \le \mu_i,\ i = 1, \ldots, m \,\}$. Непосредственная проверка показывает, что это множество выпукло. Кроме того, легко видеть, что оно содержит все векторы с положительными компонентами (надо взять $x = \widehat{x}$) и тем самым его внутренность не пуста. Наконец, $0 \notin M$, так как в противном случае нашелся бы элемент $\overline{x} \in A$ такой, что $f_i(\overline{x}) \le 0$, $i = 1, \ldots, m$, и $f_0(\overline{x}) - f_0(\widehat{x}) < 0$, в противоречие с тем, что $\widehat{x}$ — минимум.

Согласно первой теореме отделимости (см. п. 1.1.3) найдется такой ненулевой функционал, т. е. вектор $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$, что

$$\sum_{i=0}^{m} \lambda_i \mu_i \ge 0 \quad \forall\, \mu = (\mu_0, \mu_1, \ldots, \mu_m)^T \in M. \tag{i}$$

Пусть $\delta > 0$. Подставляя в (i) векторы $(1, \delta, \ldots, \delta)^T, \ldots, (\delta, \ldots, \delta, 1)^T$, а затем устремляя δ к нулю, получаем, что $\lambda_i \ge 0$, $i = 0, 1, \ldots, m$, т. е. доказано утверждение (b) теоремы.

Теперь подставим в (i) векторы $(\delta, \ldots, \delta, f_i(\widehat{x}), \delta, \ldots, \delta)^T$, $i = 1, \ldots, m$ (они принадлежат M, надо взять $x = \widehat{x}$) и снова устремляя δ к нулю, получим, что $\lambda_i f_i(\widehat{x}) \ge 0$. Но $\lambda_i f_i(\widehat{x}) \le 0$, так как $\lambda_i \ge 0$, а $f_i(\widehat{x}) \le 0$ и поэтому $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m$, что доказывает утверждение (c).

Пусть $x \in A$. Ясно, что $(f_0(x) - f_0(\widehat{x}) + \delta, f_1(x), \ldots, f_m(x))^T \in M$. Подставляя этот вектор в (i), приходим (в пределе при $\delta \to 0$) к неравенству $\sum_{i=0}^{m} \lambda_i f_i(x) \ge \lambda_0 f_0(\widehat{x})$. Добавляя справа нулевые слагаемые $\lambda_i f_i(\widehat{x})$, $i = 1, \ldots, m$, получаем, что $\mathcal{L}(x, \overline{\lambda}) \ge \mathcal{L}(\widehat{x}, \overline{\lambda})$ и (a) доказано.

Докажем второе утверждение теоремы. Пусть x — допустимый вектор в задаче (P_4). Тогда, используя это обстоятельство вместе с (b), затем (a) и (c), будем иметь

$$\lambda_0 f_0(x) \ge \lambda_0 f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) = \mathcal{L}(x, \overline{\lambda}) \ge \mathcal{L}(\widehat{x}, \overline{\lambda}) =$$

$$= \lambda_0 f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) = \lambda_0 f_0(\widehat{x}).$$

Деля на λ_0, получаем требуемое.

Докажем последнее утверждение теоремы. Если $\lambda_0 = 0$, то ненулевой множитель Лагранжа находится среди остальных и поэтому (с учетом (c)) $\mathcal{L}(\overline{x}, \overline{\lambda}) = \sum_{i=1}^{m} \lambda_i f_i(\overline{x}) < 0 = \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) = \mathcal{L}(\widehat{x}, \overline{\lambda})$, что противоречит (a). □

§ 2.5. Гладко–аппроксимативно–выпуклые задачи

Здесь рассматривается достаточно широкий класс задач, который, с одной стороны, охватывает задачи, рассмотренные в предыдущих параграфах, а с другой — задачи, которые будут рассматриваться ниже, в частности, задачи оптимального управления.

2.5.1. Постановка задачи. Пусть X и Y — нормированные пространства, $\mathcal{U}$ — топологическое пространство, V — открытое подмножество X, $f_i\colon V \times \mathcal{U} \to \mathbb{R}$, $i = 0, 1, \ldots, m$, и $F\colon V \times \mathcal{U} \to Y$. Рассмотрим задачу

$$f_0(x, u) \to \min, \quad f_i(x, u) = 0, \quad i = 1, \ldots, m, \quad F(x, u) = 0. \qquad (P)$$

Пара $(x, u) \in V \times \mathcal{U}$ называется *допустимой в задаче* (P), если $f_i(x, u) = 0$, $i = 1, \ldots, m$, и $F(x, u) = 0$. Допустимая пара $(\widehat{x}, \widehat{u})$ называется *сильным* (*локальным*) *минимумом в задаче* (P), если найдется такое $\varepsilon > 0$, что для любой допустимой пары (x, u), для которой $\|x - \widehat{x}\|_X < \varepsilon$ выполняется неравенство $f_0(x, u) \geq f_0(\widehat{x}, \widehat{u})$.

Гладкая аппроксимативная выпуклость задачи (P) означает, попросту говоря, что соответствующие отображения “гладкие” по переменной x и “почти выпуклые” по переменной u. Приведем точные определения.

Пусть X, Z — нормированные пространства, $\mathcal{U}$ — топологическое пространство, V — окрестность точки $\widehat{x} \in X$ и $\widehat{u} \in \mathcal{U}$. Скажем, что отображение $\Phi\colon V \times \mathcal{U} \to Z$ *гладко–аппроксимативно–выпукло в точке* $(\widehat{x}, \widehat{u})$, если оно

- (a) *аппроксимативно-выпукло*, т. е. для любой пары (u, v) элементов из $\mathcal{U}$, любых $\alpha \in [0, 1]$ и $\delta > 0$ найдется такой элемент $M(\alpha, \delta, u, v) \in \mathcal{U}$ (называемый *миксом* u *и* v), что $M(0, \delta, u, v) = u$, отображение $\alpha \mapsto M(\alpha, \delta, u, v)$ непрерывно на $[0, 1]$ равномерно по δ и $\Phi(x, M(\alpha, \delta, u, v)) \to (1 - \alpha)\Phi(x, u) + \alpha\Phi(x, v)$ при $\delta \to 0$ равномерно по $\alpha \in [0, 1]$ и $x \in V$;
- (b) *гладкое в* $(\widehat{x}, \widehat{u})$, т. е. непрерывно в некоторой окрестности $(\widehat{x}, \widehat{u})$, дифференцируемо по x в точке $(\widehat{x}, \widehat{u})$ и для любого $\varepsilon > 0$ найдутся такие окрестности $V(\varepsilon)$ и $U(\varepsilon)$ точек $\widehat{x}$ и $\widehat{u}$, что для всех $x, x' \in V(\varepsilon)$ и $u \in U(\varepsilon)$ справедливо неравенство
$$\|\Phi(x, u) - \Phi(x', u) - \Phi_x(\widehat{x}, \widehat{u})(x - x')\|_Z \leq \varepsilon\|x - x'\|_X.$$

Свяжем с задачей (P) функцию Лагранжа

$$\mathcal{L}((x, u), \overline{\lambda}) = \sum_{i=0}^{m} \lambda_i f_i(x, u) + \langle y^*, F(x, u)\rangle,$$

где $\overline{\lambda} = (\lambda_0, \lambda_1, \ldots, \lambda_m, y^*) \in (\mathbb{R}^{m+1})^* \times Y^*$ — набор множителей Лагранжа.

2.5.2. Правило множителей Лагранжа.

Теорема. *Пусть пара $(\widehat{x},\widehat{u})$ доставляет сильный минимум в задаче (P). Тогда, если X и Y — банаховы пространства, отображение $(f_0,f_1,\dots,f_m,F)\colon V\times\mathcal{U}\to\mathbb{R}^{m+1}\times Y$ гладко-аппроксимативно-выпукло в точке $(\widehat{x},\widehat{u})$ и $\operatorname{Im}F_x(\widehat{x},\widehat{u})=Y$, то найдутся множители Лагранжа $\lambda_i\ge 0$, $i=0,1,\dots,m$, и y^*, не равные нулю одновременно, что выполняется условие стационарности по x:*

$$\mathcal{L}_x((\widehat{x},\widehat{u}),\overline{\lambda})=0 \;\;\Leftrightarrow\;\; \sum_{i=0}^{m} f_{ix}(\widehat{x},\widehat{u})+(F_x(\widehat{x},\widehat{u}))^*y^*=0$$

и условие минимума по u:

$$\min_{u\in\mathcal{U}}\mathcal{L}((\widehat{x},u),\overline{\lambda})=\mathcal{L}((\widehat{x},\widehat{u}),\overline{\lambda}) \;\;\Leftrightarrow$$
$$\sum_{i=0}^{m} f_i(\widehat{x},u)+\langle y^*,F(\widehat{x},u)\rangle\ge\sum_{i=0}^{m} f_i(\widehat{x},\widehat{u})+\langle y^*,F(\widehat{x},\widehat{u})\rangle,\quad \forall\, u\in\mathcal{U}.$$

Доказательство. Мы приведем доказательство этой теоремы для случая, когда $m=0$, т. е. когда отсутствуют ограничения $f_i(x,u)=0$, $i=1,\dots,m$. Предварительно докажем следующее утверждение.

Лемма (о выпуклении)**.** *Если $(\widehat{x},\widehat{u})$ — сильный минимум в задаче (P), то для любого $u\in\mathcal{U}$ пара $(\widehat{x},0)$ — локальный минимум в задаче*

$$g_0(x,\alpha)=(1-\alpha)f_0(x,\widehat{u})+\alpha f_0(x,u)\to\min,$$
$$G(x,\alpha)=(1-\alpha)F(x,\widehat{u})+\alpha F(x,u)=0,\quad \alpha\ge 0. \qquad (P_u)$$

Доказательство. Пусть $u\in\mathcal{U}$ такое, что $(\widehat{x},0)$ не является локальным минимумом в (P_u). Докажем, что тогда $(\widehat{x},\widehat{u})$ — не сильный минимум в задаче (P). Для этого зафиксируем произвольную окрестность W точки $\widehat{x}$ и покажем, что найдется допустимая в задаче (P) пара (x,u) такая, что $x\in W$ и $f_0(x,u)<f_0(\widehat{x},\widehat{u})$.

Отображение F удовлетворяет теореме о неявной функции. Пусть соответствующие окрестности $V_0\subset V$ точки $\widehat{x}$, U_0 точки $\widehat{u}$, отображение φ и константа $K>0$ из этой теоремы. Уменьшая V_0, можно считать, что $V_0\subset W$.

Так как $(\widehat{x},0)$ — не локальный минимум, то в любой окрестности $(\widehat{x},0)$ найдется допустимая в (P_u) точка $(\overline{x},\overline{\alpha})$ такая, что $g_0(\overline{x},\overline{\alpha})<g_0(\widehat{x},0)=f_0(\widehat{x},\widehat{u})$. Пусть $\overline{x}\in V_0$, а $\overline{\alpha}\ge 0$ столь мало, что $\overline{\alpha}\le 1$ и (в силу свойства микса) $M(\overline{\alpha},\delta)=M(\overline{\alpha},\delta,\widehat{u},u)\in U_0$ для всех $\delta>0$. Тогда по теореме о неявной функции

$$F(\varphi(\overline{x},M(\overline{\alpha},\delta)),M(\overline{\alpha},\delta))=0 \qquad (i)$$

и

$$\|\varphi(\overline{x},M(\overline{\alpha},\delta))-\overline{x}\|_X\le K\|F(\overline{x},M(\overline{\alpha},\delta))\|_Y. \qquad (ii)$$

Из свойства гладкости f_0 следует существование таких окрестностей $V(1)$ и $U(1)$ точек $\widehat{x}$ и $\widehat{u}$, что для любых $x, x' \in V(1)$ и $u \in U(1)$ справедливо неравенство

$$|f_0(x,u) - f_0(x',u) - f_{0x}(\widehat{x},\widehat{u})(x-x')| \le \|x-x'\|_X.$$

Снова уменьшая V_0 и U_0, можно считать, что $V_0 \subset V(1)$, $U_0 \subset U(1)$ и тогда из последнего неравенства следует, что

$$|f_0(\varphi(\overline{x}, M(\overline{\alpha},\delta)), M(\overline{\alpha},\delta)) - f_0(\overline{x}, M(\overline{\alpha},\delta))| \le \\ \le 2\|f_{0x}(\widehat{x},\widehat{u})\| \|\varphi(\overline{x}, M(\overline{\alpha},\delta)) - \overline{x}\|_X. \quad (iii)$$

Поскольку $g_0(\overline{x},\overline{\alpha}) < f_0(\widehat{x},\widehat{u})$, то $g_0(\overline{x},\overline{\alpha}) < f_0(\widehat{x},\widehat{u}) - \varepsilon$ и для некоторого $\varepsilon > 0$. В силу аппроксимативной выпуклости отображения $\Phi = (f_0, F)$ в точке $(\widehat{x},\widehat{u})$ найдется такое $\overline{\delta} > 0$, что

$$\|\Phi(\overline{x}, M(\overline{\alpha},\overline{\delta})) - (1-\overline{\alpha})\Phi(\overline{x},\widehat{u}) - \overline{\alpha}\Phi(\overline{x},u)\|_{\mathbb{R}\times Y} \le \frac{\varepsilon}{2CK+1}, \quad (iv)$$

где $C = \|f_{0x}(\widehat{x},\widehat{u})\|$. Так как пара $(\overline{x},\overline{\alpha})$ допустима в задаче (P_u), то из (iv) следует, что

$$\|F(\overline{x}, M(\overline{\alpha},\overline{\delta}))\|_Y \le \frac{\varepsilon}{2CK+1}. \quad (v)$$

По построению пара $(\varphi(\overline{x}, M(\overline{\alpha},\delta)), M(\overline{\alpha},\delta))$ допустима в задаче (P) в силу (i) и $\varphi(\overline{x}, M(\overline{\alpha},\delta)) \in V_0 \subset W$. С другой стороны, используя последовательно (iii), (iv), (ii) и (v), будем иметь

$$f_0(\varphi(\overline{x}, M(\overline{\alpha},\delta), M(\overline{\alpha},\delta)) \le f_0(\overline{x}, M(\overline{\alpha},\delta)) + \\ + 2C\|\varphi(\overline{x}, M(\overline{\alpha},\delta)) - \overline{x}\|_X \le (1-\overline{\alpha})f_0(\overline{x},\widehat{u}) + \overline{\alpha}f_0(\overline{x},u) + \frac{\varepsilon}{2CK+1} + \\ + 2CK\frac{\varepsilon}{2CK+1} = g_0(\overline{x},\overline{\alpha}) + \varepsilon < f_0(\widehat{x},\widehat{u}) - \varepsilon + \varepsilon = f_0(\widehat{x},\widehat{u})$$

в противоречие с тем, что $(\widehat{x},\widehat{u})$ — сильный минимум в задаче (P). $\square$

Перейдем непосредственно к доказательству теоремы. Фиксируем $u \in \mathcal{U}$. Задача (P_u) — гладкая задача с ограничениями типа равенств и неравенств. Ее функция Лагранжа имеет вид

$$\mathcal{L}(x,\alpha,\lambda_0,\lambda_1,y^*) = \lambda_0((1-\alpha)f_0(x,\widehat{u}) + \alpha f_0(x,u)) - \lambda_1\alpha + \\ + \langle y^*, (1-\alpha)F(x,\widehat{u}) + \alpha F(x,u)\rangle.$$

Если $(\widehat{x},0)$ — решение задачи (P_u), то согласно Правилу множителей Лагранжа (см. п. 2.2.2) найдутся такие множители Лагранжа $\lambda_i = \lambda_i(u) \ge 0$, $i = 0,1$, $y^* = y^*(u)$, не равные одновременно нулю, что

$$\mathcal{L}_x(\widehat{x},0,\lambda_0,\lambda_1,y^*) = 0 \ \Leftrightarrow\ \lambda_0 f_{0x}(\widehat{x},\widehat{u}) + (F_x(\widehat{x},\widehat{u}))^* y^* = 0, \quad (i)$$

и

$$\mathcal{L}_\alpha(\widehat{x},0,\lambda_0,\lambda_1,y^*)=0 \;\Leftrightarrow\; -\lambda_0 f_0(\widehat{x},\widehat{u})+\lambda_0 f_0(\widehat{x},u)-\lambda_1- \\ -\langle y^*,F(\widehat{x},\widehat{u})\rangle+\langle y^*,F(\widehat{x},u)\rangle=0, \quad (ii)$$

откуда, учитывая, что $\lambda_1\geq 0$, следует неравенство

$$\lambda_0 f_0(\widehat{x},u)+\langle y^*,F(\widehat{x},u)\rangle\geq \lambda_0 f_0(\widehat{x},\widehat{u})+\langle y^*,F(\widehat{x},\widehat{u})\rangle. \qquad (iii)$$

Заметим теперь, что $\lambda_0\neq 0$, ибо в противном случае из (i) последовало бы, что $y^*=0$ (оператор $(F_x(\widehat{x},\widehat{u}))^*$ инъективен, так как оператор $F_x(\widehat{x},\widehat{u})$ сюръективен, см. п. 1.1.1), а тогда из (ii) вытекало бы, что и $\lambda_1=0$, что невозможно. Итак, $\lambda_0\neq 0$ и можно считать, что $\lambda_0=1$. Тогда из (i) следует (снова в силу взаимной однозначности $(F_x(\widehat{x},\widehat{u}))^*$), что y^* не зависит от u и значит, соотношения (i) и (iii) и есть утверждения теоремы. □

Глава 3. Условия экстремума в задачах вариационного исчисления и оптимального управления

§ 3.1. Задача Больца — гладкая задача без ограничений

3.1.1. Постановка задачи. Пусть $[t_0,t_1]$ — отрезок числовой прямой, $L\colon \mathbb{R}^3\to\mathbb{R}$ — непрерывная функция переменных $t,x,\dot{x}$ и $l\colon\mathbb{R}^2\to\mathbb{R}$ — непрерывная функция переменных x_0 и x_1.

Задача

$$\mathcal{B}(x(\cdot))=\int_{t_0}^{t_1} L(t,x(t),\dot{x}(t))\,dt+l(x(t_0),x(t_1))\to\min, \\ x(\cdot)\in C^1([t_0,t_1]) \quad (P_5)$$

называется *задачей Больца вариационного исчисления.*

Согласно §2.1 это задача без ограничений, где $X=C^1([t_0,t_1])$ (см. п. 1.1.2).

Локальный минимум в этой задаче называется *слабым локальным минимумом.*

Необходимые условия минимума в гладкой задаче без ограничений даются теоремой Ферма (см. §2.1). Следующее утверждение есть расшифровка этой теоремы для данного случая.

3.1.2. Необходимые условия первого порядка в задаче Больца. Далее, если фиксирована функция $\widehat{x}(\cdot)$, то для сокращения записи используем обозначения: $\widehat{L}_x(t)=L_x(t,\widehat{x}(t),\dot{\widehat{x}}(t))$ и аналогично для частной производной L по $\dot{x}$, а также $\widehat{l}_{x_i}=l_{x_i}(\widehat{x}(t_0),\widehat{x}(t_1))$, $i=0,1$.

Теорема. *Пусть $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (P_5). Тогда, если функция L непрерывна вместе со своими частными производными по x и $\dot{x}$ в $\mathbb{R}^3$, а функция l дифференцируема в точке $(\widehat{x}(t_0), \widehat{x}(t_1))$, то для всех $t \in [t_0, t_1]$ выполнено уравнение Эйлера:*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0$$

и условия трансверсальности:

$$\widehat{L}_{\dot{x}}(t_i) = (-1)^i \widehat{l}_{x_i}, \quad i = 0, 1,$$

Доказательство. Покажем, что минимизируемый функционал в (P_5) дифференцируем в точке $\widehat{x}(\cdot)$ и для любого $x(\cdot) \in C^1([t_0, t_1])$

$$\langle \mathcal{B}'(\widehat{x}(\cdot)), x(\cdot)\rangle = \int_{t_0}^{t_1} (\widehat{L}_x(t)x(t) + \widehat{L}_{\dot{x}}(t)\dot{x}(t))\,dt + \widehat{l}_{x_0}x(t_0) + \widehat{l}_{x_1}x(t_1). \quad (i)$$

Действительно, первое слагаемое этого функционала есть суперпозиция трех отображений: линейного непрерывного отображения $x(\cdot) \mapsto (x(\cdot), \dot{x}(\cdot))$ из $C^1([t_0, t_1])$ в $C^1([t_0, t_1]) \times C([t_0, t_1])$, отображения $(x(\cdot), u(\cdot)) \mapsto L(\cdot, x(\cdot), u(\cdot))$ из $C^1([t_0, t_1]) \times C([t_0, t_1])$ в $C([t_0, t_1]$ и линейного непрерывного функционала $g(\cdot) \mapsto \int_{t_0}^{t_1} g(t)\,dt$ на $C([t_0, t_1])$. Первое и третье отображения дифференцируемы, второе (в силу условий теоремы) также дифференцируемо в $\widehat{x}(\cdot)$ (см. п. 1.2.2 примеры 2, 3 и следствие). Тогда по теореме о суперпозиции дифференцируемых отображений (п. 1.2.1) получаем, что интеграл в (i) есть производная первого слагаемого в функционале $\mathcal{B}$ в точке $\widehat{x}(\cdot)$.

Второе слагаемое есть суперпозиция линейного непрерывного отображения $x(\cdot) \mapsto (x(t_0), x(t_1))$ из $C^1([t_0, t_1])$ в $\mathbb{R}^2$ и функции $(x_0, x_1) \mapsto l(x_0, x_1)$ на $\mathbb{R}^2$, которая по условию дифференцируема в $(\widehat{x}(t_0), \widehat{x}(t_1))$. Тогда снова по теореме о суперпозиции дифференцируемых отображений получаем, что производная второго слагаемого в функционале $\mathcal{B}$ в точке $\widehat{x}(\cdot)$ равна сумме второго и третьего слагаемых (i). Соотношение (i) доказано.

Так как $\widehat{x}(\cdot)$ — локальный минимум, то по теореме Ферма $\mathcal{B}'(\widehat{x}(\cdot)) = 0$. Покажем, что это равносильно утверждению теоремы. Для этого представим функционал $\mathcal{B}'(\widehat{x}(\cdot))$ в каноническом виде (см. п. 1.1.2, формула (R_1)). Интегрируя по частям, будем иметь

$$\int_{t_0}^{t_1} \widehat{L}_x(t)x(t)\,dt = x(t_0)\int_{t_0}^{t_1} \widehat{L}_x(t)\,dt + \int_{t_0}^{t_1} \dot{x}(t)\left(\int_t^{t_1} \widehat{L}_x(\tau)\,d\tau\right)dt =$$
$$= x(t_0)\int_{t_0}^{t_1} \widehat{L}_x(t)\,dt + \int_{t_0}^{t_1} \dot{x}(t)\,d\left(\int_{t_0}^{t} ds \int_s^{t_1} \widehat{L}_x(\tau)\,d\tau\right).$$

Запишем еще

$$\int_{t_0}^{t_1} \widehat{L}_{\dot{x}}(t)\dot{x}(t)\,dt = \int_{t_0}^{t_1} \dot{x}(t)\,d\left(\int_{t_0}^{t} \widehat{L}_{\dot{x}}(\tau)\,d\tau\right)$$

и $x(t_1) = x(t_0) + \int_{t_0}^{t_1} \dot{x}(t)\,dt$. Тогда

$$\langle \mathcal{B}'(\widehat{x}(\cdot)), x(\cdot)\rangle = \langle a, x(t_0)\rangle + \int_{t_0}^{t_1} \dot{x}(t)\,dv(t), \qquad (ii)$$

где $a = \int_{t_0}^{t_1} \widehat{L}_x(t)\,dt + \widehat{l}_{x_0} + \widehat{l}_{x_1}$ и

$$v(t) = \int_{t_0}^{t} \widehat{L}_{\dot{x}}(\tau)\,d\tau + \int_{t_0}^{t} ds \int_{s}^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{x_1}(t - t_0), \quad t \in [t_0, t_1].$$

Элементарно проверяется, что $v(\cdot)$ — непрерывная функция ограниченной вариации, равная нулю в t_0.

В силу единственности представления (ii) (см. пример 3 п. 1.1.2) $\mathcal{B}'(\widehat{x}(\cdot)) = 0$ тогда и только тогда, когда $a = 0$ и $v(\cdot) = 0$. Дифференцируя тождество $v(t) \equiv 0$, получаем, что для любого $t \in [t_0, t_1]$

$$\widehat{L}_{\dot{x}}(t) + \int_{t}^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{x_1} = 0. \qquad (iii)$$

Дифференцируя теперь это тождество, получаем уравнение Эйлера. Условия трансверсальности легко следуют из соотношения (iii) при $t = t_0$ и t_1 и соотношения $a = 0$. $\square$

§ 3.2. Простейшая задача вариационного исчисления

3.2.1. Постановка задачи. Пусть $[t_0, t_1]$ — отрезок числовой прямой, $L\colon \mathbb{R}^3 \to \mathbb{R}$ — непрерывная функция переменных $t, x, \dot{x}$ и $x_0, x_1 \in \mathbb{R}$.

Задача

$$J(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\,dt \to \min, \quad x(t_0) = x_1, \quad x(t_1) = x_1,$$
$$x(\cdot) \in C^1([t_0, t_1]) \quad (P_6)$$

называется *простейшей задачей вариационного исчисления.*

Как и в задаче Больца, локальный минимум в этой задаче называется *слабым локальным минимумом.*

Полагая $X = C^1([t_0, t_1])$, $f\colon X \to \mathbb{R}$, $f(x(\cdot)) = J(x(\cdot))$ и $F\colon X \to \mathbb{R}^2$, $F(x(\cdot)) = (x(t_0) - x_0, x(t_1) - x_1)$, получаем, что задача (P_6) есть частный случай общей задачи (P_2) из §2.2, результатами которого мы и воспользуемся для исследования задачи

Локальный минимум в задаче (P_6) называется *слабым локальным минимумом.*

3.2.2. Необходимые условия первого порядка в простейшей задаче — уравнение Эйлера. Как и раньше, если фиксирована функция $\widehat{x}(\cdot)$, то для сокращения записи используем обозначения: $\widehat{L}_x(t) = L_x(t, \widehat{x}(t), \dot{\widehat{x}}(t))$ и $\widehat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$.

Теорема. *Пусть* $\widehat{x}(\cdot)$ *— слабый локальный минимум в задаче* (P_6). *Тогда, если функция* L *непрерывна вместе со своими частными производными по* x *и* $\dot{x}$ *в* $\mathbb{R}^3$, *то для всех* $t \in [t_0, t_1]$ *выполнено уравнение Эйлера:*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0.$$

Доказательство. Покажем, что уравнение Эйлера есть просто расшифровка правила множителей Лагранжа — теоремы из п. 2.2.2. Проверим, что все условия этой теоремы выполнены. Действительно, при доказательстве теоремы из п. 3.1.2 показано, что функционал J дифференцируем в точке $\widehat{x}(\cdot)$. Отображение F аффинно и поэтому дифференцируемо (см. п. 1.2.2) и, наконец, подпространство $\operatorname{Im} F'(\widehat{x}(\cdot))$ в $\mathbb{R}^2$ замкнуто, так как конечномерно. Запишем функцию Лагранжа задачи (P_6), учитывая, что в нашем случае $Y = \mathbb{R}^2$ и поэтому $y^* = (\lambda_1, \lambda_2) \in (\mathbb{R}^2)^*$:

$$\mathcal{L}(x(\cdot), \overline{\lambda}) = \lambda_0 f(x) + \langle y^*, F(x) \rangle = \lambda_0 \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt +$$
$$+ \lambda_1(x(t_0) - x_0) + \lambda_2(x(t_1) - x_1), \quad (i)$$

где $\overline{\lambda} = (\lambda_0, \lambda_1, \lambda_2)$. Согласно правилу множителей Лагранжа найдутся ненулевой вектор $\overline{\lambda}$ такой, что $\mathcal{L}_{x(\cdot)}(\widehat{x}(\cdot), \overline{\lambda}) = 0$. Заметим теперь, функция Лагранжа $\mathcal{L}$ имеет вид функционала Больца $\mathcal{B}$ (п. 3.1.1), где надо заменить L на $\lambda_0 L$, а $l(x(t_0), x(t_1))$ на $\lambda_1(x(t_0) - x_0) + \lambda_2(x(t_1) - x_1)$. Тогда равенство $\mathcal{L}_{x(\cdot)}(\widehat{x}(\cdot), \overline{\lambda}) = 0$, по уже доказанному, равносильно тому, что выполняются соотношения

$$-\frac{d}{dt}\lambda_0\widehat{L}_{\dot{x}}(t) + \lambda_0\widehat{L}_x(t) = 0, \quad (ii)$$

$\lambda_0\widehat{L}_{\dot{x}}(t_0) = \lambda_1$ и $\lambda_0\widehat{L}_{\dot{x}}(t_1) = -\lambda_2$. Из последних двух следует, что $\lambda_0 \neq 0$ (иначе $\overline{\lambda} = 0$), а тогда деля (ii) на λ_0, получаем утверждение теоремы. $\square$

3.2.3. Необходимые условия второго порядка в простейшей задаче. Условия Лежандра и Якоби. Мы воспользуемся первой теоремой п. 2.2.3. Для этого сначала приведем условия, гарантирующие дважды дифференцируемость минимизируемого функционала в задаче (P_6). Как и для первых производных, для краткости, пишем $\widehat{L}_{xx}(t) = L_{xx}(t, \widehat{x}(t), \dot{\widehat{x}}(t))$ и аналогично для других производных.

Предложение 2. *Пусть функция* $L\colon \mathbb{R}^3 \to \mathbb{R}$ *и ее частные производные по* x *и* $\dot{x}$ *до второго порядка включительно существуют и непрерывны на* $\mathbb{R}^3$. *Тогда функционал* $J\colon C^1([t_0, t_1]) \to \mathbb{R}$ *задачи* (P_6) *дважды непрерывно дифференцируем в каждой точке*

$\widehat{x}(\cdot) \in C^1([t_0,t_1])$ *и для любых* $x_i(\cdot) \in C^1([t_0,t_1])$, $i=1,2$, *справедливо представление*

$$J''(\widehat{x}(\cdot))[x_1(\cdot),x_2(\cdot)] = \int_{t_0}^{t_1} (\widehat{L}_{xx}(t)x_1(t)x_2(t) + \widehat{L}_{\dot{x}x}(t)x_1(t)\dot{x}_2(t) + \\ + \widehat{L}_{x\dot{x}}(t)\dot{x}_1(t)x_2(t) + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{x}_1(t)\dot{x}_2(t))\,dt.$$

Доказательство этого факта не приводим — оно получается на том же пути, что и для первой производной функционала J, но с небольшими техническими усложнениями.

Теорема (Необходимые условия второго порядка для простейшей задачи)**.** *Пусть* $\widehat{x}(\cdot)$ *— слабый локальный минимум в задаче* (P_6). *Тогда, если функция* L *и ее частные производные по* x *и* $\dot{x}$ *до второго порядка включительно существуют и непрерывны на* $\mathbb{R}^3$, *то для всех* $t \in [t_0,t_1]$ *выполнено уравнение Эйлера*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0$$

и для всех $h(\cdot) \in C^1([t_0,t_1])$ *таких, что* $h(t_0) = h(t_1) = 0$ *справедливо неравенство*

$$\int_{t_0}^{t_1} (\widehat{L}_{xx}(t)h^2(t) + 2\widehat{L}_{\dot{x}x}(t)h(t)\dot{h}(t) + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t))\,dt \geq 0. \qquad (3.1)$$

Доказательство. В п. 3.2.1 показано как сводится задача (P_6) к общей задаче (P_2). Проверим, что выполнены условия первой теоремы п. 2.2.3. Из предложения 2 следует, что функционал $f = J$ дважды дифференцируем в точке $\widehat{x}(\cdot)$. Отображение $F\colon x(\cdot) \to (x(t_0)-x_0, x(t_1)-x_1)$ аффинно и его производная в любой точке действует по правилу: $F'(\widehat{x}(\cdot))[x(\cdot)] = (x(t_0), x(t_1))$ (см. п. 1.2.2). Так как для любой пары $(a,b) \in \mathbb{R}^2$, очевидно, найдется функция $x(\cdot) \in C^1([t_0,t_1])$ такая, что $x(t_0) = a$ и $x(t_1) = b$, то оператор $F'(\widehat{x}(\cdot))$ сюрьективен. Тогда согласно теореме п. 2.2.3 найдется функционал $y^* \in (\mathbb{R}^2)^*$, т. е. пара (λ_1,λ_2), такая, что $\mathcal{L}_{x(\cdot)}(\widehat{x}(\cdot),1,\lambda_1,\lambda_2) = 0$. Это равносильно, как показано в предыдущей теореме, уравнению Эйлера. Условие (3) теоремы п. 2.2.3 равносильно соотношению (3.1) доказываемой теоремы, поскольку $\operatorname{Ker} F'(\widehat{x}(\cdot)) = \{\, h(\cdot) \in C^1([t_0,t_1]) \mid h(t_0) = h(t_1) = 0 \,\}$, вторая производная F равна нулю (см. п. 1.2.2) и тем самым вторая производная функции Лагранжа совпадает со второй производной функционала J, подсчитанной в предложении 2 (надо еще учесть, что $\widehat{L}_{x\dot{x}}(\cdot) = \widehat{L}_{\dot{x}x}(\cdot)$ в силу непрерывности этих производных). $\square$

Дальнейшие исследования связаны с проверкой условия (3.1) — знакоопределенности квадратичной формы $J''(\widehat{x}(\cdot))$. Это условие, очевидно, равносильно тому что функция $\widehat{h}(\cdot) = 0$ является глобальным минимумом задачи

$$\mathcal{K}(h(\cdot)) = J''(\widehat{x}(\cdot))[h(\cdot), h(\cdot)] \to \min, \quad h(\cdot) \in C^1([t_0, t_1]),$$
$$h(t_0) = h(t_1) = 0, \quad (3.2)$$

которая есть частный случай простейшей задачи вариационного исчисления и потому $\widehat{h}(\cdot)$ (а также любое другое решение (3.2)) должно удовлетворять уравнению Эйлера

$$-\frac{d}{dt}\left[\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t) + \widehat{L}_{\dot{x}x}(t)h(t)\right] + \widehat{L}_{\dot{x}x}(t)\dot{h}(t) + \widehat{L}_{xx}(t)h(t) = 0. \quad (3.3)$$

Это дифференциальное уравнение (уравнение Эйлера для задачи (3.2)) называется *уравнением Якоби* для задачи (P_6).

Введем некоторые определения. Пусть $\widehat{x}(\cdot)$ — *экстремаль* задачи (P_6), т. е. $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера для этой задачи.

(0.1) Говорят, что на функции $\widehat{x}(\cdot)$ выполнено *условие Лежандра*, если $\widehat{L}_{\dot{x}\dot{x}}(t) \geq 0$ для всех $t \in [t_0, t_1]$ и *усиленное условие Лежандра*, если $\widehat{L}_{\dot{x}\dot{x}}(t) > 0$ для всех $t \in [t_0, t_1]$.

(0.2) Пусть на $\widehat{x}(\cdot)$ выполнено усиленное условие Лежандра. Точка $\tau \in (t_0, t_1]$ называется *сопряженной точкой* к точке t_0, если существует нетривиальное решение $h(\cdot)$ уравнения Якоби такое, что $h(t_0) = h(\tau) = 0$.

(0.3) Говорят, что на функции $\widehat{x}(\cdot)$ выполнено *условие Якоби*, если в интервале (t_0, t_1) нет точек, сопряженных с t_0 и *усиленное условие Якоби*, если в полуинтервале $(t_0, t_1]$ нет точек, сопряженных с t_0.

Заметим, что если выполнено усиленное условие Лежандра, то уравнение Якоби (3.3) равносильно следующей системе линейных уравнений

$$\begin{cases} \dot{x} = -\dfrac{\widehat{L}_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}x + \dfrac{1}{\widehat{L}_{\dot{x}\dot{x}}(t)}y \\[2ex] \dot{y} = \left(\widehat{L}_{xx}(t) - \dfrac{\widehat{L}^2_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}\right)x + \dfrac{\widehat{L}_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}y. \end{cases} \quad (3.4)$$

Действительно, обозначая $y = \widehat{L}_{\dot{x}\dot{x}}(t)\dot{x} + \widehat{L}_{\dot{x}x}(t)x$, после несложных преобразований, приходим к данной системе. Для такой системы справедлива глобальная (т. е. на всем отрезке $[t_0, t_1]$) теорема существования и единственности решения при любых начальных условиях.

Рассмотрим следующую экстремальную задачу

$$\mathcal{K}(h(\cdot)) \to \min, \quad h(\cdot) \in W_2^1([t_0, t_1]), \quad h(t_0) = h(t_1) = 0, \quad (3.2')$$

Эта задача отличается от задачи (3.2) тем, что пространство $C^1([t_0,t_1])$ заменено на пространство $W_2^1([t_0,t_1])$, в котором (как хорошо известно) $C^1([t_0,t_1])$ плотно. Отсюда легко извлечь, что если нулевая функция является решением задачи (3.2), то она является и решением данной задачи (докажите это).

Предложение 3. *Решение* $\widehat{h}(\cdot)$ *задачи* $(3.2')$ *удовлетворяет уравнению Якоби* (3.3).

Доказательство проведите сами, воспользовавшись следующими указаниями: 1) расшифруйте правило множителей Лагранжа из п. 2.2.2 для задачи $(3.2')$; 2) представьте, интегрируя по частям, производную по x функции Лагранжа в каноническом виде в пространстве $W_2^1([t_0,t_1])$ и воспользуйтесь теоремой из примера 5 п. 1.1.2 (см. аналогичные рассуждения в теореме 3.2.1).

Предложение 4. *Пусть интегрант* L *имеет непрерывные частные производные по* x *и* $\dot{x}$ *до второго порядка включительно,* $\widehat{x}(\cdot)\in C^1([t_0,t_1])$ *и* $\mathcal{K}(h(\cdot)) = J''(\widehat{x}(\cdot))[h(\cdot),h(\cdot)] \geq 0$ *для всех* $h(\cdot)\in C^1([t_0,t_1])$ *таких, что* $h(t_0)=h(t_1)=0$. *Тогда* $\widehat{L}_{\dot{x}\dot{x}}(t) = L_{\dot{x}\dot{x}}(t,\widehat{x}(t),\dot{\widehat{x}}(t)) \geq 0$ *для любого* $t\in[t_0,t_1]$.

Доказательство. Согласно условию и из вышесказанного следует, что в задаче $(3.2')$ минимум достигается в нуле. Предположим, что утверждение предложения неверно. Тогда существует точка $\tau\in[t_0,t_1]$ такая, что $\widehat{L}_{\dot{x}\dot{x}}(\tau)<0$. В силу непрерывности $\widehat{L}_{\dot{x}\dot{x}}(\cdot)$ можно считать, что $\tau\in(t_0,t_1)$ и, очевидно, найдутся числа $\varepsilon_0>0$ и v такие, что $[\tau-\varepsilon_0/2,\tau+\varepsilon_0/2]\subset(t_0,t_1)$ и $\widehat{L}_{\dot{x}\dot{x}}(t)v^2<-1$ для всех $t\in[\tau-\varepsilon/2,\tau+\varepsilon/2]$ и $0<\varepsilon\leq\varepsilon_0$. Для каждого $0<\varepsilon\leq\varepsilon_0$ положим

$$\widetilde{h}_\varepsilon(t)=\begin{cases} v(t-\tau+\varepsilon/2), & t\in[\tau-\varepsilon/2,\tau];\\ -v(t-\tau-\varepsilon/2), & t\in[\tau,\tau+\varepsilon/2];\\ 0, & |t-\tau|>\varepsilon/2.\end{cases}$$

Ясно, что $\widetilde{h}_\varepsilon(\cdot)\in W_2^1([t_0,t_1])$, $\widetilde{h}_\varepsilon(t_0)=\widetilde{h}_\varepsilon(t_1)=0$ и

$$\dot{\widetilde{h}}_\varepsilon(t)=\begin{cases} v, & t\in(\tau-\varepsilon/2,\tau);\\ -v, & t\in(\tau,\tau+\varepsilon/2);\\ 0, & |t-\tau|>\varepsilon/2.\end{cases}$$

Поскольку

$$\int_{t_0}^{t_1}\widehat{L}_{\dot{x}\dot{x}}(t)(\dot{\widetilde{h}}_\varepsilon(t))^2\,dt=\int_{\tau-\varepsilon/2}^{\tau+\varepsilon/2}\widehat{L}_{\dot{x}\dot{x}}(t)v^2\,dt<-\varepsilon,$$

$\max_{t\in[t_0,t_1]}\widetilde{h}_\varepsilon(t)=\tau\varepsilon/2$ и

$$\int_{t_0}^{t_1}\widehat{L}_{\dot{x}x}(t)\widetilde{h}_\varepsilon(t)\dot{\widetilde{h}}_\varepsilon(t)\,dt\leq\frac{|v|\tau\varepsilon}{2}\int_{\tau-\varepsilon/2}^{\tau+\varepsilon/2}\widehat{L}_{\dot{x}x}(t)\,dt\leq C\varepsilon^2,$$

где $C = (|v|\tau/2)\max_{t\in[t_0,t_1]}|\widehat{L}_{\dot{x}x}(t)|$, то непосредственный подсчет показывает, что $\mathcal{K}(\widetilde{h}_\varepsilon(\cdot)) < C_1\varepsilon^2-\varepsilon$, т. е. $\mathcal{K}(\widetilde{h}_\varepsilon(\cdot)) < 0$ для достаточно малого ε в противоречие с тем, что решением задачи $(3.2')$ служит нулевая функция. □

Предложение 5. *Пусть выполнены предположения предложения 4 и $\widehat{L}_{\dot{x}\dot{x}}(t) > 0$ для любого $t \in [t_0, t_1]$. Тогда на интервале (t_0, t_1) нет точек сопряженных с t_0.*

Доказательство. Доказываем от противного. Пусть $\tau \in (t_0, t_1)$ и $\overline{h}(\cdot)$ — нетривиальное решение уравнения Якоби такое, что $\overline{h}(t_0) = \overline{h}(\tau) = 0$. Положим

$$\widetilde{h}(t) = \begin{cases} \overline{h}(t), & t \in [t_0, \tau]; \\ 0, & t \in [\tau, t_1]. \end{cases}$$

Вычислим значение функционала $\mathcal{K}$ на этой функции. Интегрируя по частям, будем иметь

$$\begin{aligned} \mathcal{K}(\widetilde{h}(\cdot)) &= \int_{t_0}^{\tau} \Big(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}^2(t) + \widehat{L}_{xx}(t)\overline{h}^2(t) + 2\widehat{L}_{\dot{x}x}(t)\overline{h}(t)\dot{\overline{h}}(t)\Big)\,dt = \\ &= \int_{t_0}^{\tau}\Big(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}(t) + \widehat{L}_{\dot{x}x}(t)\overline{h}(t)\Big)\dot{\overline{h}}(t)\,dt + \int_{t_0}^{\tau}\Big(\widehat{L}_{x\dot{x}}(t)\dot{\overline{h}}(t) + \\ &\quad + \widehat{L}_{xx}(t)\overline{h}(t)\Big)\overline{h}(t)\,dt = \int_{t_0}^{\tau}\Big(-\frac{d}{dt}\Big(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}(t) + \\ &\quad + \widehat{L}_{\dot{x}x}(t)\overline{h}(t)\Big) + \widehat{L}_{x\dot{x}}(t)\dot{\overline{h}}(t) + \widehat{L}_{xx}(t)\overline{h}(t)\Big)\overline{h}(t)\,dt. \end{aligned}$$

Поскольку $\overline{h}(\cdot)$ удовлетворяет уравнению Якоби, то отсюда следует, что $\mathcal{K}(\widetilde{h}(\cdot)) = 0$. Таким образом, в силу условия предложения функция $\widetilde{h}(\cdot)$ является решением задачи $(3.2')$. Поэтому в силу предложения 3 эта функция удовлетворяет уравнению Якоби (3.3). Покажем, что, на самом деле, это невозможно, а именно, что в этом случае непрерывная функция $p(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)\dot{\widetilde{h}}(\cdot) + \widehat{L}_{\dot{x}x}(\cdot)\widetilde{h}(\cdot)$ оказывается разрывной в точке τ. В самом деле, $\widetilde{h}(\cdot) = 0$ на $[\tau, t_1]$ и поэтому $p(\tau+0) = 0$. С другой стороны, $p(\tau-0) = \widehat{L}_{\dot{x}\dot{x}}(\tau)\dot{\widetilde{h}}(\tau-0) = \widehat{L}_{\dot{x}\dot{x}}(\tau)\dot{\overline{h}}(\tau)$. Заметим теперь, что $\dot{\overline{h}}(\tau) \neq 0$, ибо в противном случае функция $\overline{h}(\cdot)$ удовлетворяла бы уравнению Якоби с начальными условиями $\overline{h}(\tau) = \dot{\overline{h}}(\tau) = 0$, но тогда, в силу отмеченной выше равносильности уравнения Якоби (3.3) и линейной системы (3.4) функция $\overline{h}(\cdot)$ была бы тождественным нулем, что по предположению не так. Далее $\widehat{L}_{\dot{x}\dot{x}}(\tau) > 0$, так как выполнено усиленное условие Лежандра и значит, $p(\tau-0) \neq 0$, т. е. функция $p(\cdot)$ разрывна в точке τ. Полученное противоречие доказывает теорему. □

Как итог предыдущий рассмотрений получаем следующее утверждение.

Теорема (Необходимые условия Лежандра и Якоби)**.** *Пусть* $\widehat{x}(\cdot)$ *доставляет слабы локальный минимум в задаче* (P_6). *Тогда, если функция* L *и ее частные производные по* x *и* $\dot{x}$ *до второго порядка включительно существуют и непрерывны на* $\mathbb{R}^3$, *то* $\widehat{x}(\cdot)$ *удовлетворяет уравнению Эйлера, на* $\widehat{x}(\cdot)$ *выполнено условие Лежандра, и если на* $\widehat{x}(\cdot)$ *выполнено усиленное условие Лежандра, то выполнено и условие Якоби.*

3.2.4. Достаточные условия второго порядка в простейшей задаче.

Предложение 6. *Пусть функция* $L\colon \mathbb{R}^3 \to \mathbb{R}$ *и ее частные производные по* x *и* $\dot{x}$ *до второго порядка включительно существуют и непрерывны на* $\mathbb{R}^3$. *Тогда для минимизируемого функционала* J *в простейшей задаче вариационного исчисления* (P) *справедливо разложение второго порядка в любой точке* $\widehat{x}(\cdot) \in C^1([t_0,t_1])$ *относительно пары* $(C^1([t_0,t_1]), W_2^1([t_0,t_1]))$.

Доказательство. Согласно определению в п. 2.2.3 надо проверить, что пространство $C^1([t_0,t_1])$ непрерывно вложено в $W_2^1([t_0,t_1])$, что функционал J дважды дифференцируем в $\widehat{x}(\cdot)$, его вторая производная $J''(\widehat{x}(\cdot))$ непрерывна, как билинейная форма, на $W_2^1([t_0,t_1]) \times W_2^1([t_0,t_1]$ и для всех $x(\cdot) \in C^1([t_0,t_1])$ справедливо представление

$$J(\widehat{x}(\cdot)+x(\cdot)) = J(\widehat{x}(\cdot)) + J'(\widehat{x}(\cdot))x(\cdot) + \frac{1}{2}J''(\widehat{x}(\cdot))[x(\cdot),x(\cdot)] + r(x(\cdot)), \qquad (i)$$

где $r(x(\cdot))/\|x(\cdot)\|^2_{W_2^1([t_0,t_1])} \to 0$ при $x(\cdot) \to 0$ в $C^1([t_0,t_1])$.

Непрерывность вложения $C^1([t_0,t_1])$ в $W_2^1([t_0,t_1])$ проверяется элементарно. Доказательство формулы (i) не приводим — оно получается на том же пути, что и формулы для первой и второй производной функционала J. $\square$

Теорема (Достаточные условия второго порядка для простейшей задачи в терминах квадратичной формы)**.** *Пусть функция* $L\colon \mathbb{R}^3 \to \mathbb{R}$ *и ее частные производные по* x *и* $\dot{x}$ *до второго порядка включительно существуют и непрерывны на* $\mathbb{R}^3$. *Тогда, если* $\widehat{x}(\cdot)$ *— допустимая экстремаль в задаче* (P_6) *и существует такое число* $\alpha > 0$, *что*

$$\int_{t_0}^{t_1} (\widehat{L}_{xx}(t)h^2(t) + 2\widehat{L}_{\dot{x}x}(t)h(t)\dot{h}(t) + \widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t))\,dt \geq \alpha\|h(\cdot)\|_{W_2^1([t_0,t_1])} \qquad (3.5)$$

для всех $h(\cdot) \in C^1([t_0,t_1])$, *для которых* $h(t_0) = h(t_1) = 0$, *то* $\widehat{x}(\cdot)$ *— слабый локальный минимум в задаче* (P_6).

Доказательство. Как уже было показано в п. 3.2.1 задача (P_6) есть частный случай общей задачи (P_2), а именно, $X = C^1([t_0, t_1])$, $f\colon X \to \mathbb{R}$, $f(x(\cdot)) = J(x(\cdot))$ и $F\colon X \to \mathbb{R}^2$, $F(x(\cdot)) = (x(t_0) - x_0, x(t_1) - x_1)$. Покажем теперь, что утверждения теоремы есть расшифровка условий теоремы об обобщенных достаточных условиях п. 2.2.3. Проверим, что выполнены все требования последней теоремы. Как было уже отмечено, элементарно проверяется непрерывность вложения $C^1([t_0, t_1])$ в $W_2^1([t_0, t_1])$. Из предложения 6 следует, что для функционала J справедливо разложение второго порядка в точке $\widehat{x}(\cdot)$ относительно пары $(C^1([t_0, t_1]), W_2^1([t_0, t_1]))$. Так как отображение F аффинно, то для него такое разложение выполняется очевидным образом. Пространство $\operatorname{Im} F'(\widehat{x})$ замкнуто, так как конечномерно. По условию $\widehat{x}(\cdot)$ — экстремаль задачи (P_6), т. е. удовлетворяет уравнению Эйлера, которое, как показано в доказательстве теоремы п. 3.2.2, равносильно тому, что $\mathcal{L}_{x(\cdot)}(\widehat{x}(\cdot), \overline{\lambda}) = 0$, где $\mathcal{L}$ — функция Лагранжа задачи (P_6) (см. (i) в п. 3.2.2). Условие

$$\mathcal{L}_{xx}(\widehat{x}, 1, y^*)[h, h] \geq \alpha \|h\|^2_{X_1}, \ \ \forall\, h \in \operatorname{Ker} F'(\widehat{x})$$

теоремы п. 2.2.3 есть условие доказываемой теоремы, где $X_1 = W_2^1([t_0, t_1])$, так как $\operatorname{Ker} F'(\widehat{x}(\cdot)) = \{\, h(\cdot) \in C^1([t_0, t_1]) \mid h(t_0) = h(t_1) = 0 \,\}$, вторая производная F равна нулю и тем самым вторая производная функции Лагранжа совпадает со второй производной функционала J, подсчитанной в предложении 2. Итак, все условия теоремы об обобщенных достаточных условиях п. 2.2.3 выполнены и, следовательно, $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (P_6). □

Дальнейшие исследования достаточных условий минимума для простейшей задачи вариационного исчисления связаны с проверкой условия (3.5). Пусть $A(\cdot)$, $B(\cdot)$ и $C(\cdot)$ — непрерывные функции на $[t_0, t_1]$. Рассмотрим квадратичный функционал

$$Q(h(\cdot)) = \int_{t_0}^{t_1} (A(t)\dot{h}^2(t) + 2C(t)\dot{h}(t)h(t) + B(t)h^2(t))\, dt$$

на функциях $h(\cdot) \in C^1([t_0, t_1])$ таких, что $h(t_0) = h(t_1) = 0$. Выпишем для него формально уравнение Эйлера

$$-\frac{d}{dt}\Big[A(t)\dot{h}(t) + C(t)h(t)\Big] + C(t)\dot{h}(t) + B(t)h(t) = 0. \tag{3.6}$$

Если $A(t) > 0$ для всех $t \in [t_0, t_1]$, то уравнение (3.6) равносильно линейной системе уравнений (это уже было отмечено выше, см. доказанную равносильность (3.4) и (3.5)) и значит, для любых начальных данных существует единственное решение уравнения (3.6), определенное на всем отрезке $[t_0, t_1]$.

Предложение 7. *Пусть $A(t) > 0$ для всех $t \in [t_0, t_1]$ и решение уравнения* (3.6) *с начальными данными $h(t_0) = 0$, $\dot{h}(t_0) = 1$ не* *обращается в ноль на полуинтервале* $(t_0, t_1]$. *Тогда* $Q(h(\cdot)) > 0$ *для всех ненулевых* $h(\cdot) \in C^1([t_0, t_1])$ *таких, что* $h(t_0) = h(t_1) = 0$.

Доказательство. Справедливо следующее тождество

$$\int_{t_0}^{t_1} (A(t)\dot h^2(t) + 2C(t)\dot h(t)h(t) + B(t)h^2(t))\,dt =$$
$$= \int_{t_0}^{t_1} A(t)\left(\dot h(t) - \frac{\dot w(t)}{w(t)}h(t)\right)^2 dt, \quad (3.7)$$

где $w(\cdot)$ — решение уравнения (3.6), не обращающееся в ноль на отрезке $[t_0, t_1]$. Такое решение существует в силу условий предложения и теоремы о непрерывной зависимости решения от начальных данных. Кроме того, функция $(\dot h(\cdot) - (\dot w(\cdot)/w(\cdot))h(\cdot)$ не есть тождественный ноль и поэтому из (3.7) следует нужное утверждение.

Доказательство самого тождества (3.7) не приводим — оно носит технический характер. □

Следствие. *Пусть выполнены условия предложения* 7. *Тогда существует такое* $\alpha > 0$, *что всех* $h(\cdot) \in C^1([t_0, t_1])$, *для которых* $h(t_0) = h(t_1) = 0$, *справедливо неравенство*

$$Q(h(\cdot)) \geq \alpha \|h(\cdot)\|^2_{W_2^1([t_0,t_1])}. \quad (3.8)$$

Доказательство. Так как функция $A(\cdot)$ непрерывна и положительна на $[t_0, t_1]$, то для любого $0 < \alpha < \min_{t\in[t_0,t_1]} A(t)$ функция $A_\alpha(\cdot) = A(\cdot) - \alpha$ также будет положительна на $[t_0, t_1]$. Для каждого такого α положим

$$Q_\alpha(h(\cdot)) = \int_{t_0}^{t_1} (A_\alpha(t)\dot h^2(t) + 2C(t)\dot h(t)h(t) + B(t)h^2(t))\,dt$$

и запишем соответствующее уравнение Эйлера

$$-\frac{d}{dt}\left[A_\alpha(t)\dot h(t) + C(t)h(t)\right] + C(t)\dot h(t) + B(t)h(t) = 0. \quad (3.6\alpha)$$

По условию решение уравнения (3.6) с начальными данными $h(t_0) = 0$, $\dot h(t_0) = 1$ не обращается в ноль на полуинтервале $(t_0, t_1]$ и поэтому в силу теоремы о непрерывной зависимости решения от параметров, решение уравнения (3.6α) с теми же начальными данными также не будет иметь нулей на $(t_0, t_1]$ для достаточно малых $\alpha > 0$. Для таких α, в силу предложения 7, имеем $Q_\alpha(h(\cdot)) > 0$, если $h(\cdot) \in C^1([t_0, t_1])$, $h(\cdot) \neq 0$ и $h(t_0) = h(t_1) = 0$. Следовательно, $Q(h(\cdot)) = Q_\alpha(h(\cdot)) + \alpha\int_{t_0}^{t_1} \dot h^2(t)\,dt \geq \alpha\|h(\cdot)\|^2_{W_2^1([t_0,t_1])}$. □

Теперь все готово для получения классических достаточных условий минимума для простейшей задачи вариационного исчисления (P_6). Справедлива следующая

Теорема (Достаточные условия минимума в форме усиленных условий Лежандра и Якоби)**.** *Пусть функция $L\colon \mathbb{R}^3 \to \mathbb{R}$ и ее частные производные по x и $\dot{x}$ до второго порядка включительно существуют и непрерывны на $\mathbb{R}^3$. Тогда, если на допустимой экстремали $\widehat{x}(\cdot)$ задачи (P_6) выполнены усиленные условия Лежандра и Якоби, то $\widehat{x}(\cdot)$ — слабый локальный минимум в этой задаче.*

Доказательство. Легко проверить, что если выполнено усиленное условие Лежандра, то усиленное условие Якоби равносильно тому, что решение уравнения Якоби с начальными условиями $h(t_0) = 0$, $\dot{h}(t_0) = 1$ не обращается в ноль на полуинтервале $(t_0, t_1]$ и поэтому выполнены условия предложения 7 для квадратичной формы $Q(h(\cdot)) = J''(\widehat{x}(\cdot))[h(\cdot), h(\cdot)]$. Тогда из следствия этого предложения и теоремы о достаточных условиях минимума в терминах квадратичной формы вытекает, что $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (P_6). □

§ 3.3. Задача Лагранжа — гладкая задача с ограничениями типа равенств

3.3.1. Постановка задачи. Пусть $[t_0, t_1]$ — отрезок числовой прямой, функция $L_0\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}$, отображение $\varphi\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}^n$ (переменных $t \in \mathbb{R}$, $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, $u = (u_1, \dots, u_r)^T \in \mathbb{R}^r$) и функции $l_i\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$ (переменных ξ_0 и ξ_1) непрерывны на своей области определения. Задача

$$J(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} L_0(t, x(t), u(t))\, dt \to \min, \quad \dot{x} = \varphi(t, x, u),$$
$$l_i(x(t_0), x(t_1)) = 0, \quad i = 1, \dots, m \quad (P_7)$$

называется *задачей Лагранжа вариационного исчисления.*

Уточним постановку. Пара $(x(\cdot), u(\cdot))$ называется *допустимой в задаче* (P_7), если $(x(\cdot), u(\cdot)) \in C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$, $\dot{x}(t) = \varphi(t, x(t), u(t))$ для всех $t \in [t_0, t_1]$ и $l_i(x(t_0), x(t_1)) = 0$, $1 \le i \le m$.

Допустимая пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ называется *слабым локальным минимумом в задаче* (P_7), если существует такое $\varepsilon > 0$, что для любой допустимой пары $(x(\cdot), u(\cdot))$, для которой $\|(x(\cdot), u(\cdot)) - (\widehat{x}(\cdot), \widehat{u}(\cdot))\|_{C^1([t_0,t_1],\mathbb{R}^n)\times C([t_0,t_1],\mathbb{R}^r)} < \varepsilon$ выполнено неравенство $J(x(\cdot), u(\cdot)) \ge J(\widehat{x}(\cdot), \widehat{u}(\cdot))$.

Задача (P_7) является частным случаем задачи (P_2) — гладкой задачи с ограничениями типа равенств. Действительно, положим $X = C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$, $Y = C([t_0, t_1], \mathbb{R}^n) \times \mathbb{R}^m$, $f(x) = J(x(\cdot), u(\cdot))$ и $F(x)(t) = (\dot{x}(t) - \varphi(t, x(t), u(t)), l_1(x(t_0), x(t_1)), \dots, l_m(x(t_0), x(t_1)))$ для любого $t \in [t_0, t_1]$.

Функция Лагранжа задачи (P_7) (в силу структуры сопряженного пространства к произведению пространств, теоремы Рисса и вида сопряженного к $\mathbb{R}^m$, см. п. 1.1.2) имеет вид

$$\mathcal{L}(x,\lambda_0,y^*) = \mathcal{L}((x(\cdot),u(\cdot)),\lambda_0,\mu(\cdot),\lambda_1,\ldots,\lambda_m) =$$
$$= \lambda_0\int_{t_0}^{t_1} L_0(t,x(t),u(t))\,dt + \int_{t_0}^{t_1}(\dot x(t)-\varphi(t,x(t),u(t))\,d\mu(t) +$$
$$+\sum_{i=1}^m \lambda_i l_i(x(t_0),x(t_1)).$$

3.3.2. Уравнение Эйлера–Лагранжа — необходимое условие минимума первого порядка. Пусть $p\in(\mathbb{R}^n)^*$. Положим $L(t,x,\dot x,u,\lambda_0,p) = \lambda_0 L_0(t,x,u) - \langle p,\dot x-\varphi(t,x,u)\rangle$ и $l(\lambda,\xi_0,\xi_1) = \sum_{i=1}^m \lambda_i l_i(\xi_0,\xi_1)$, где $\lambda=(\lambda_1,\ldots,\lambda_m)$. Если фиксированы функции $\widehat x(\cdot)\in C^1([t_0,t_1],\mathbb{R}^n)$, $\widehat u(\cdot)\in C([t_0,t_1],\mathbb{R}^r)$, $\widehat p(\cdot)\colon[t_0,t_1]\to(\mathbb{R}^n)^*$, число $\widehat\lambda_0$ и вектор $\widehat\lambda=(\widehat\lambda_1,\ldots,\widehat\lambda_m)$, то, как и раньше, для краткости записи, пишем $\widehat L_x(t) = L_x(t,\widehat x(t),\dot{\widehat x}(t),\widehat u(t),\widehat\lambda_0,\widehat p(t))$ и аналогично для частной производной по u и частных производных функции l и отображения φ.

Теорема (Необходимые условия минимума первого порядка для задачи (P_7))**.** *Пусть пара $(\widehat x(\cdot),\widehat u(\cdot))$ доставляет слабый локальный минимум в задаче (P_7). Тогда, если функция L_0 и отображение φ непрерывны вместе со своими частными производными по x и u на $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r$, а функции l_i, $1\le i\le m$, непрерывно дифференцируемы в точке $(\widehat x(t_0),\widehat x(t_1))$, то найдутся, не равные одновременно нулю, число $\widehat\lambda_0$, вектор $\widehat\lambda=(\widehat\lambda_1,\ldots,\widehat\lambda_m)$ и функция $p(\cdot)\in C^1([t_0,t_1],(\mathbb{R}^n)^*)$ такие, что для всех $t\in[t_0,t_1]$ выполнено уравнение Эйлера по x:*

$$-\frac{d}{dt}\widehat L_{\dot x}(t) + \widehat L_x(t) = 0 \ \Leftrightarrow\ -\dot p = p\widehat\varphi_x(t) - \lambda_0\widehat L_{0x}(t),$$

уравнение Эйлера по u:

$$\widehat L_u(t) = 0 \ \Leftrightarrow\ p(t)\widehat\varphi_u(t) = \lambda_0\widehat L_{0u}(t)$$

и условия трансверсальности:

$$\widehat L_{\dot x}(t_i) = (-1)^i\widehat l_{\xi_i},\quad i=0,1.$$

Доказательство. Покажем, что утверждения теоремы есть просто расшифровка утверждения общей теоремы п. 2.2.2. Проверим, что все ее условия выполнены. Функция L_0 и отображение φ суть обобщенные операторы Немыцкого (см. п. 1.2.2), которые в силу условий теоремы непрерывно дифференцируемы в точке $(\widehat x(\cdot),\widehat u(\cdot))$ (см. там же). В частности, функционал $f = J$

дифференцируем в точке $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ (как суперпозиция дифференцируемой функции с линейным функционалом), а отображение F непрерывно дифференцируемо в этой точке. Тогда согласно следствию из п. 1.2.1 отображение F строго дифференцируемо в точке $(\widehat{x}(\cdot), \widehat{u}(\cdot))$. Отображение F можно записать в виде $F = (F_1, F_2)$, где $F_1\colon C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r) \to C([t_0, t_1], \mathbb{R}^n)$, $F_1(x(\cdot), u(\cdot))(t) = \dot{x}(t) - \varphi(t, x(t), u(t))$, $t \in [t_0, t_1]$, и производная F_1 в точке $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ действует по правилу $F_1'(\widehat{x}(\cdot), \widehat{u}(\cdot))[x(\cdot), u(\cdot)](t) = \dot{x}(t) - \widehat{\varphi}_x(t)x(t) - \widehat{\varphi}_u(t)u(t)$ (см. п. 1.2.2). Это сюръективное отображение. Действительно, пусть $y(\cdot) \in C([t_0, t_1], \mathbb{R}^n)$. Согласно теореме существования для линейной системы уравнений существует функция $\overline{x}(\cdot)$ на $[t_0, t_1]$ (удовлетворяющая, скажем, условиям $x(t_0) = x_0$) такая, что $\dot{\overline{x}}(t) = \widehat{\varphi}_x(t)\overline{x}(t) + y(t)$ для всех $t \in [t_0, t_1]$. Это означает, что $F_1'(\widehat{x}(\cdot), \widehat{u}(\cdot))[\overline{x}(\cdot), 0](t) = y(t)$, $t \in [t_0, t_1]$. Производная отображения F_2 действует в $\mathbb{R}^m$ и поэтому по лемме о замкнутости образа (п. 1.1.4) образ оператора $F'(\widehat{x}(\cdot), \widehat{u}(\cdot))$ замкнут. Условия теоремы п. 2.2.2 выполнены.

Согласно этой теореме найдутся множители Лагранжа $\widehat{\lambda}_0$, $\widehat{\lambda} = (\widehat{\lambda}_1, \ldots, \widehat{\lambda}_m)$ и $\widehat{\mu}(\cdot)$, не равные одновременно нулю, такие, что в точке $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ выполняется условие $\mathcal{L}'((\widehat{x}(\cdot), \widehat{u}(\cdot)), \widehat{\lambda}_0, \widehat{\mu}(\cdot), \widehat{\lambda}_1, \ldots, \widehat{\lambda}_m) = 0$, или, равносильно, частные производные функции Лагранжа по $x(\cdot)$ и по $u(\cdot)$ в точке $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ равняются нулю. Производные L_0 и φ уже были подсчитаны и поэтому мы имеем

$$\langle \mathcal{L}_{x(\cdot)}((\widehat{x}(\cdot), \widehat{u}(\cdot)), \widehat{\lambda}_0, \widehat{\mu}(\cdot), \widehat{\lambda}_1, \ldots, \widehat{\lambda}_m), x(\cdot)\rangle = \widehat{\lambda}_0 \int_{t_0}^{t_1} \langle \widehat{L}_{0x}(t), x(t)\rangle\, dt +$$

$$+ \int_{t_0}^{t_1} (\dot{x}(t) - \widehat{\varphi}_x(t)x(t))\, d\mu(t) + \langle \widehat{l}_{\xi_0}, x(t_0)\rangle + \langle \widehat{l}_{\xi_1}, x(t_1)\rangle = 0 \quad (i)$$

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

$$\langle \mathcal{L}_{u(\cdot)}((\widehat{x}(\cdot), \widehat{u}(\cdot)), \widehat{\lambda}_0, \widehat{\mu}(\cdot), \widehat{\lambda}_1, \ldots, \widehat{\lambda}_m), u(\cdot)\rangle = \widehat{\lambda}_0 \int_{t_0}^{t_1} \langle \widehat{L}_{0u}(t), u(t)\rangle\, dt -$$

$$- \int_{t_0}^{t_1} \widehat{\varphi}_u(t)u(t)\, d\mu(t) = 0 \quad (ii)$$

для любого $u(\cdot) \in C([t_0, t_1], \mathbb{R}^r)$.

Покажем, что эти соотношения — равенства нулю функционалов на $C^1([t_0, t_1], \mathbb{R}^n)$ и $C([t_0, t_1], \mathbb{R}^r)$ — равносильны утверждениям теоремы. Для этого представим эти функционалы в каноническом виде (см. п. 1.1.2). Обозначим для этого через x_1^* функционал, равный сумме первого, третьего и четвертого слагаемых в правой части первого равенства в (i). Тогда повторяя, фактически, рассуждения п. 3.1.2. с заменой обычных функций на вектор-функции, получаем, что для всех $x(\cdot) \in C^1([t_0,t_1],\mathbb{R}^n)$ справедливо представление

$$\langle x_1^*, x(\cdot)\rangle = \langle a_1, x(t_0)\rangle + \int_{t_0}^{t_1} \dot{x}(t)\, d\nu_1(t), \tag{iii}$$

где $a_1 = \widehat{\lambda}_0 \int_{t_0}^{t_1} \widehat{L}_{0x}(t)\,dt + \widehat{l}_{\xi_0} + \widehat{l}_{\xi_1}$ и $\nu_1(t) = \widehat{\lambda}_0 \int_{t_0}^{t} d\tau \int_{\tau}^{t_1} \widehat{L}_{0x}(s)\,ds + (t-t_0)\widehat{l}_{\xi_1}$. Далее преобразуем второе слагаемое в правой части первого равенства в (i), используя равенство $x(t) = x(t_0) + \int_{t_0}^{t} \dot{x}(\tau)\,d\tau$ и меняя порядок интегрирования в двойном интеграле (чтобы не усложнять запись, считаем, что $n=1$)

$$\int_{t_0}^{t_1} (\dot{x}(t) - \widehat{\varphi}_x(t)x(t))\,d\mu(t) = \int_{t_0}^{t_1} \dot{x}(t)\,d\mu(t) - x(t_0)\int_{t_0}^{t_1} \widehat{\varphi}_x(t)\,d\mu(t) -$$
$$- \int_{t_0}^{t_1} \widehat{\varphi}_x(t) \left(\int_{t_0}^{t} \dot{x}(\tau)\,d\tau\right) d\mu(t) = \int_{t_0}^{t_1} \dot{x}(t)\,d\mu(t) - x(t_0)\int_{t_0}^{t_1} \widehat{\varphi}_x(t)\,d\mu(t) -$$
$$- \int_{t_0}^{t_1} \dot{x}(\tau) \left(\int_{\tau}^{t_1} \widehat{\varphi}_x(t)\,d\mu(t)\right) d\tau = -x(t_0)\int_{t_0}^{t_1} \widehat{\varphi}_x(t)\,d\mu(t) +$$
$$+ \int_{t_0}^{t_1} \dot{x}(t) d\left(\mu(t) - \int_{t_0}^{t}\left(\int_{\tau}^{t_1} \widehat{\varphi}_x(s)\,d\mu(s)\right) d\tau\right).$$

Объединяя это с (iii), соотношение (i) теперь записывается в виде (для краткости левую часть (i) обозначаем $\langle \widehat{\mathcal{L}}_{x(\cdot)}, x(\cdot)\rangle$)

$$\langle \widehat{\mathcal{L}}_{x(\cdot)}, x(\cdot)\rangle = \langle a, x(t_0)\rangle + \int_{t_0}^{t_1} \dot{x}(t)\,d\nu(t) = 0, \tag{iv}$$

где $a = a_1 - \int_{t_0}^{t_1} \widehat{\varphi}_x(t)\,d\mu(t)$ и

$$\nu(t) = \nu_1(t) + \mu(t) - \int_{t_0}^{t}\left(\int_{\tau}^{t_1} \widehat{\varphi}_x(s)\,d\mu(s)\right) d\tau. \tag{v}$$

В силу единственности канонического представления функционала (см. п. 1.1.2) равенство (iv) возможно тогда и только тогда, когда $a=0$ и $\nu(\cdot)=0$. Так как, очевидно, $\nu_1(\cdot) \in C^1([t_0,t_1],\mathbb{R}^n)$, то из равенства $\nu(\cdot)=0$ и (v) следует, что $\mu(\cdot) \in C^1([t_0,t_1],\mathbb{R}^n)$. Положим $p(\cdot) = \dot{\mu}(\cdot)$ (тогда $d\mu(t) = p(t)dt$). Дифференцируя тождество $\nu(t) \equiv 0$, $t \in [t_0,t_1]$, получим

$$\widehat{\lambda}_0 \int_{t}^{t_1} \widehat{L}_{0x}(s)\,ds + \widehat{l}_{\xi_1} + p(t) - \int_{t}^{t_1} p(s)\widehat{\varphi}_x(s)\,ds = 0 \tag{vi}$$

для любого $t \in [t_0,t_1]$. Дифференцируя теперь это тождество, приходим к уравнению Эйлера по x.

Условия трансверсальности следуют из (vi) при $t=t_0$ и t_1 и соотношения $a=0$.

Осталось доказать уравнение Эйлера по u. Рассуждая аналогично предыдущему, записываем равенство (ii) в виде

$$\langle \widehat{\mathcal{L}}_{u(\cdot)}, u(\cdot)\rangle = \int_{t_0}^{t_1} u(t)\, d\eta(t) = 0,$$

где (учитывая, что $d\mu(t) = p(t)dt$)

$$\eta(t) = \int_{t_0}^{t} (\widehat{\lambda}_0 \widehat{L}_{0x}(\tau)\, d\tau - p(\tau)\widehat{\varphi}_x(\tau))\, d\tau, \quad \forall\, t \in [t_0, t_1].$$

Снова в силу единственности представления необходимо $\eta(\cdot) = 0$. Тогда дифференцируя тождество $\eta(t) \equiv 0$, получаем уравнение Эйлера по u. $\square$

§ 3.4. Задача оптимального управления — гладко-аппроксимативно-выпуклая задача

3.4.1. Постановка задачи. Пусть $[t_0, t_1]$ — отрезок числовой прямой, U — непустое подмножество $\mathbb{R}^r$, функция $f\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}$ и отображение $\varphi\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r \to \mathbb{R}^n$ (переменных $t \in \mathbb{R}$, $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ и $u = (u_1, \dots, u_r)^T \in \mathbb{R}^r$) непрерывны на своей области определения и $x_i \in \mathbb{R}^n$, $i = 0, 1$. Задача

$$J(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} L_0(t, x(t), u(t))\, dt \to \min, \quad \dot{x} = \varphi(t, x, u),$$
$$u(t) \in U, \quad x(t_0) = x_0, \quad x(t_1) = x_1 \quad (P_8)$$

называется *задачей оптимального управления* с закрепленными концами. Переменную $x(\cdot)$ часто называют фазовой переменной, а $u(\cdot)$ — управлением.

Уточним постановку. Пусть $PC^1([t_0, t_1], \mathbb{R}^n)$ — совокупность всех кусочно-непрерывно дифференцируемых, а $PC([t_0, t_1], \mathbb{R}^r)$ — кусочно-непрерывных функций на $[t_0, t_1]$ со значениями соответственно в $\mathbb{R}^n$ и $\mathbb{R}^r$. Пара $(x(\cdot), u(\cdot)) \in PC^1([t_0, t_1], \mathbb{R}^n) \times PC([t_0, t_1], \mathbb{R}^r)$ называется *допустимой в задаче* (P_8), если включение $u(t) \in U$ и равенство $\dot{x}(t) = \varphi(t, x(t), u(t))$ выполняются для всех $t \in [t_0, t_1]$, где функция $u(\cdot)$ непрерывна и $x(t_i) = x_i$, $i = 0, 1$.

Допустимая пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ называется *сильным (локальным) минимумом в задаче* (P_8), если существует такое $\varepsilon > 0$, что для любой допустимой пары $(x(\cdot), u(\cdot))$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)} < \varepsilon$ выполнено неравенство $J(x(\cdot), u(\cdot)) \geq J(\widehat{x}(\cdot), \widehat{u}(\cdot))$.

Для вывода необходимых условий минимума воспользуемся общей теорией, изложенной в § 2.5. Для этого запишем задачу (P_8) в форме задачи (P) п. 2.5.1. Пусть $X = Y = C([t_0, t_1], \mathbb{R}^n)$, $\mathcal{U} = \{\, u(\cdot) \in PC([t_0, t_1], \mathbb{R}^r) \mid u(t) \in U$ в точках непрерывности $u(\cdot)\,\}$ с метрикой, индуцированной из $L_1([t_0, t_1], \mathbb{R}^r)$ (совокупность суммируемых функций $u(\cdot)$ на $[t_0, t_1]$ со значениями в $\mathbb{R}^r$ и нормой

$\int_{t_0}^{t_1} |u(t)|\,dt)$, $f_0(x(\cdot),u(\cdot)) = J(x(\cdot),u(\cdot))$, $f_i(x(\cdot),u(\cdot)) = x_i(t_1) - x_{i1}$, $i = 1,\dots,n$ (где $x(t_j) = (x_1(t_j),\dots,x_n(t_j)$ и $x_j = (x_{ij},\dots,x_{nj})$, $j = 0,1)$,

$$F(x(\cdot),u(\cdot))(t) = x(t) - x_0 - \int_{t_0}^{t} \varphi(\tau,x(\tau),u(\tau))\,d\tau, \quad \forall\, t \in [t_0,t_1].$$

Равенство $F(x(\cdot),u(\cdot))(t) = 0$ для всех $t \in [t_0,t_1]$ равносильно тому, что $x(\cdot)$ является решением задачи Коши $\dot{x} = \varphi(t,x,u(t))$, $x(t_0) = x_0$, и, таким образом, задача (P_8) записана в форме задачи (P).

Функция Лагранжа задачи (P_8) имеет вид

$$\mathcal{L}((x(\cdot),u(\cdot)),\overline{\lambda}) = \lambda_0 \int_{t_0}^{t_1} L_0(t,x(t),u(t))\,dt + \sum_{i=1}^{n} \lambda_i(x_i(t_1) - x_{i1}) + \\ + \int_{t_0}^{t_1} \left(x(t) - x_0 - \int_{t_0}^{t} \varphi(\tau,x(\tau),u(\tau))\,d\tau \right) d\mu(t),$$

где $\overline{\lambda} = (\lambda_0,\lambda_1,\dots,\lambda_n,\mu(\cdot))$, а $\mu(\cdot)$ — функция ограниченной вариации, непрерывная справа на (t_0,t_1) и равная нулю в t_0.

Пусть $p \in (\mathbb{R}^n)^*$. Как и в задаче Лагранжа положим $L(t,x,\dot{x},u,\lambda_0,p) = \lambda_0 L_0(t,x,u) - \langle p,\dot{x} - \varphi(t,x,u)\rangle$

Функцию $H(t,x,u,\lambda_0,p) = \langle p,\varphi(t,x,u)\rangle - \lambda_0 L_0(t,x,u)$ называют *функцией Понтрягина задачи* (P_8).

Ниже пользуемся теми же соглашениями об обозначениях, о которых сказано в п. 3.3.2.

3.4.2. Принцип максимума Понтрягина — необходимые условия первого порядка.

Теорема (Принцип максимума Понтрягина)**.** *Пусть пара* $(\widehat{x}(\cdot),\widehat{u}(\cdot))$ *доставляет сильный минимум в задаче* (P_8). *Тогда, если функция* L_0 *и отображение* φ *непрерывны вместе со своими частными производными по* x, *то найдутся, не равные одновременно нулю, число* $\lambda_0 \geq 0$ *и функция* $p(\cdot) \in PC^1([t_0,t_1],(\mathbb{R}^n)^*)$ *такие, что выполнено уравнение Эйлера* $\Leftrightarrow$ *сопряженное уравнение для* $p(\cdot)$:

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) = 0 \quad \Leftrightarrow \quad \dot{p} = -H_x(t,\widehat{x}(t),\widehat{u}(t),\lambda_0,p(t))$$

и в точках непрерывности $\widehat{u}(\cdot)$ *условие минимума по* $u(\cdot)$:

$$\min_{u\in U} L(t,\widehat{x}(t),\dot{\widehat{x}}(t),u,\lambda_0,p(t)) = L(t,\widehat{x}(t),\dot{\widehat{x}}(t),\widehat{u}(t),\lambda_0,p(t))$$

или (что то же) условие максимума

$$\max_{u\in U} H(t,\widehat{x}(t),u,\lambda_0,p(t)) = H(t,\widehat{x}(t),\widehat{u}(t),\lambda_0,p(t)).$$

Из-за последнего соотношения необходимые условия в задаче оптимального управления и называют “Принципом максимума Понтрягина”.

Доказательство. Для произвольных элементов $u_1(\cdot), u_2(\cdot) \in \mathcal{U}$, чисел $\alpha \in [0,1]$ и $\delta > 0$ построим микс $M(\alpha, \delta, u_1(\cdot), u_2(\cdot))(\cdot)$ следующим образом. Разобьем отрезок $[t_0, t_1]$ на подотрезки Δ длины не больше δ ($|\Delta| \le \delta$). Затем каждый из таких подотрезков разобьем на два подподотрезка Δ_1 и Δ_2 длины соответственно $(1-\alpha)|\Delta|$ и $\alpha|\Delta|$. Положим $M(\alpha, \delta, u_1(\cdot), u_2(\cdot))(t) = u_i(t)$, если $t \in \Delta_i$, $i = 1, 2$.

Отображение $\Phi = (f_0, F)$ гладко-аппроксимативно-выпукло и при этом для любого $x(\cdot) \in C([t_0, t_1], \mathbb{R}^n)$ справедливы соотношения

$$F_{x(\cdot)}(\widehat{x}(\cdot), \widehat{u}(\cdot))[x(\cdot)](t) = x(t) - \int_{t_0}^{t} \widehat{\varphi}_x(\tau) x(\tau)\, d\tau, \quad \forall\, t \in [t_0, t_1] \qquad (i)$$

и

$$\langle \mathcal{L}_{x(\cdot)}((\widehat{x}(\cdot), \widehat{u}(\cdot)), \overline{\lambda}), x(\cdot) \rangle = \lambda_0 \int_{t_0}^{t_1} \widehat{L}_{0x}(t) x(t)\, dt + \int_{t_0}^{t_1} (x(t) -$$
$$- \int_{t_0}^{t} \widehat{\varphi}_x(\tau) x(\tau)\, d\tau)\, d\mu(t) + \sum_{i=1}^{n} \lambda_i x_i(t_1). \qquad (ii)$$

Доказательство всех этих фактов несложно, но достаточно громоздко и поэтому соответствующие выкладки мы опускаем.

Для того, чтобы воспользоваться общей теоремой осталось проверить, что $\operatorname{Im} F_{x(\cdot)}(\widehat{x}(\cdot), \widehat{u}(\cdot)) = C([t_0, t_1], \mathbb{R}^n)$. Действительно, пусть $y(\cdot) \in C([t_0, t_1], \mathbb{R}^n)$. Надо доказать, согласно (i), существование такой функции $\overline{x}(\cdot) \in C([t_0, t_1], \mathbb{R}^n)$, что $\overline{x}(t) - \int_{t_0}^{t} \widehat{\varphi}_x(\tau) \overline{x}(\tau)\, d\tau = y(t)$ для всех $t \in [t_0, t_1]$. Задача Коши для линейного уравнения $\dot{z} = \widehat{\varphi}_x(t) z + \widehat{\varphi}_x(t) y(t)$, $z(t_0) = 0$, имеет единственное решение $\overline{z}(\cdot)$, определенное на всем отрезке $[t_0, t_1]$. Это равносильно тому, что $\overline{z}(t) = \int_{t_0}^{t} \widehat{\varphi}_x(\tau)(\overline{z}(\tau) + y(\tau))\, d\tau$ для всех $t \in [t_0, t_1]$. Обозначая $\overline{x}(\cdot) = \overline{z}(\cdot) + y(\cdot)$, получаем требуемое.

Таким образом, выполнены все предположения теоремы п. 2.5.2. Покажем, что ее утверждения равносильны утверждениям доказываемой теоремы. Представим для этого функционал (ii) в каноническом виде (см. п. 1.1.2). С одной стороны,

$$\lambda_0 \int_{t_0}^{t_1} \widehat{L}_{0x}(t) x(t)\, dt = \int_{t_0}^{t_1} x(t) d \left(\int_{t_0}^{t} \lambda_0 \widehat{L}_{0x}(\tau)\, d\tau \right).$$

Далее интегрируем по частям (считая для простоты выкладок $n = 1$)

$$\int_{t_0}^{t_1} \left(\int_{t_0}^{t} \widehat{\varphi}_x(\tau) x(\tau)\, d\tau \right) d\mu(t) = \mu(t_1) \int_{t_0}^{t_1} \widehat{\varphi}_x(t) x(t)\, dt -$$
$$- \int_{t_0}^{t_1} \mu(t) \widehat{\varphi}_x(t) x(t)\, dt = \int_{t_0}^{t_1} x(t) d \left(\int_{t_0}^{t_1} \widehat{\varphi}_x(\tau)(\mu(t_1) - \mu(\tau))\, d\tau \right).$$

Наконец

$$\sum_{i=1}^{n} \lambda_i x_i(t_1) = \sum_{i=1}^{n} \int_{t_0}^{t_1} x_i(t)\, d\eta_i(t) = \int_{t_0}^{t_1} x(t)\, d\eta(t),$$

где $\eta(\cdot) = (\eta_1(\cdot), \dots, \eta_n(\cdot))$, $\eta_i(t) = 0$, если $t \in [t_0, t_1)$ и $\eta_i(t_1) = \lambda_i$, $i = 1, \dots, n$.

Теперь формулу (ii) можно записать так

$$\langle \mathcal{L}_{x(\cdot)}((\widehat{x}(\cdot), \widehat{u}(\cdot)), \overline{\lambda}), x(\cdot)\rangle = \int_{t_0}^{t_1} x(t)\, d\nu(t), \qquad (iii)$$

где

$$\nu(t) = \lambda_0 \int_{t_0}^{t_1} \widehat{L}_{0x}(\tau)\, d\tau + \mu(t) - \int_{t_0}^{t} \widehat{\varphi}_x(\tau)(\mu(t_1) - \mu(\tau))\, d\tau + \eta(t).$$

Первое утверждение теоремы п. 2.5.2 состоит в том, что функционал (iii) нулевой. В силу единственности это равносильно тому, что $\nu(\cdot) = 0$. Для любого $t \in [t_0, t_1]$ равенство $\nu(t) - \nu(t_1) = 0$ имеет вид

$$\mu(t) - \mu(t_1) - \int_{t}^{t_1} (\lambda_0 L_{0x}(\tau) - \widehat{\varphi}_x(\tau)(\mu(t_1) - \mu(\tau)))\, d\tau + \eta(t) - \lambda = 0, \quad (iv)$$

где $\lambda = (\lambda_1, \dots, \lambda_n)$.

Положим $p(t) = \mu(t) - \mu(t_1)$, если $t \in [t_0, t_1)$ и $p(t_1) = -\lambda$. Тогда из тождества $\nu(t) \equiv 0$, $t \in [t_0, t_1]$ и (iv) следует, что

$$p(t) = \int_{t}^{t_1} H_x(\tau, \widehat{x}(\tau), \widehat{u}(\tau), \lambda_0, p(\tau))\, d\tau - \lambda$$

Дифференцируя это равенство в точках непрерывности $\widehat{u}(\cdot)$, получаем первое утверждение доказываемой теоремы.

Расшифровывая, аналогично предыдущему, второе утверждение общей теоремы п. 2.5.2, приходим к соотношению

$$\int_{t_0}^{t_1} H_x(t, \widehat{x}(t), \widehat{u}(t), \lambda_0, p(t))\, dt = \max_{u(\cdot) \in \mathcal{U}} \int_{t_0}^{t_1} H_x(t, \widehat{x}(t), u(t), \lambda_0, p(t))\, dt,$$

из которого легко следует второе утверждение доказываемой теоремы. $\square$