Вариационное исчисление и оптимальное управление (мехмат МГУ, лекции 2016 г.)

Е. Р. Аваков, Г. Г. Магарил-Ильяев

Содержание

Введение

Какие причины побуждают решать задачи на максимум и минимум? Интерес к экстремальным задачам, т. е. задачам на максимум и минимум проявился уже на заре развития математики, и основными стимулами были любознательность и стремление к совершенству.

Среди наиболее ранних, точно решенных задач — так называемая *изопериметрическая задача* — задача о форме кривой заданной длины, охватывающей наибольшую площадь[1] и задача о форме поверхности заданной площади, охватывающей наибольший объем. Ответы на эти задачи для мыслителей Древней Греции были символами совершенства человеческого разума. Крупнейшие их представители: Евклид, Архимед и Аполлоний ставили и решали различные геометрические задачи на экстремум. Задача *о параллелограмме наибольшей площади, который можно вписать в треугольник* приводится в "Началах" Евклида (III в. до н. э.); задача *о шаровом сегменте максимального объема при заданной площади шаровой части поверхности этого сегмента* содержится в

[1]Ответ в ней приводил в своих сочинениях еще Аристотель (IV в. до н. э.)

сочинениях Архимеда (тоже III в. до н. э.); задача *о минимальном расстоянии от точки плоскости до эллипса и о нормалях к эллипсу из произвольной точки плоскости* была поставлена и решена Аполлонием (III–II в. до н. э.) в его знаменитых “Кониках”.

Важная причина, побуждающая интерес к исследованию экстремальных задач, связана с тем, что многие законы природы, как оказалось, подчинены *экстремальным принципам*, т. е. многие природные процессы (неким загадочным образом) являются решениями задач на максимум и минимум. Л. Эйлер по этому поводу высказался так: “В мире не происходит ничего, в чем не был бы виден смысл какого-нибудь максимума или минимума”.

Первый экстремальный принцип в естествознании выдвинул Пьер Ферма (1662). Согласно этому принципу *свет “избирает” такую траекторию, вдоль которой время, затрачиваемое им на преодоление пути от одной точки до другой, минимально.* Законы преломления света, установленные экспериментально, находятся, исходя из этого принципа, как решения экстремальных задач. Отправляясь от принципа Ферма, И. Бернулли в 1696 году решил так называемую *задачу о брахистохроне* — о кривой наибыстрейшего ската, т. е. задачу о форме кривой, соединяющей две точки в вертикальной плоскости, вдоль которой тело под действием силы тяжести без трения проходит путь от одной точки до другой за кратчайшее время (постановка такой задачи возможно была навеяна более ранними размышлениями Галилея на эту тему). Историю развития вариационного исчисления принято отсчитывать именно с этого времени — года решения задачи о брахистохроне.

Нельзя не назвать также и прагматические причины, по которым приходится решать экстремальные задачи. Людям свойственно наилучшим образом использовать ресурсы, находящиеся в их распоряжении, и потому экстремальные задач естественно возникают при управлении различными процессами, в экономике, инженерии.

Первой задачей такого рода была *аэродинамическая задача Ньютона* о форме тела вращения, испытывающем наименьшее сопротивление в некоей “редкой среде”, поставленная и решенная Ньютоном в его “Математических началах натуральной философии” (1687 г.). К числу экстремальных задач, возникающих в экономике и теории управления, относятся, например, *транспортная задача* (сформулированная в сороковые годы прошлого века) о наилучшем способе транспортировки продуктов со складов в магазины и *простейшая задача о быстродействии*, формализующая, в частности, задачу о наименьшем времени движения лифта в угольной шахте. Эта задача, также сформулированная в сороковые годы, послужила стимулом для создания *теории оптимального* *управления.* Считалось, что эта теория родилась в пятидесятые годы прошлого века, но когда она, в основном, сложилась, выяснилось, что первой задачей оптимального управления была именно аэродинамическая задача Ньютона.

Экстремальные задачи и их формализация. Мы видим, что зачастую экстремальные задачи ставятся на языке той области знаний, из которой они происходят. Для того, чтобы эти задачи исследовать математическими средствами, необходимо перевести их на математический язык, т. е. *формализовать*. Точно поставленная экстремальная задача включает в себя функционал f (со значениями, вообще говоря, в расширенной прямой $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$) вместе со своей областью определения X и множество ограничений $C \subset X$. Формализованная экстремальная задача записывается так

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

и заключается в нахождении таких точек $x \in C$, в которых функционал f достигает своего минимума (максимума) на C. Такие точки называются *глобальными* или *абсолютными минимумами* (*максимумами*) в задаче (P) или ее *решениями*. Если нас интересуют и точки минимума и точки максимума, то вместо $\min(\max)$ пишем extr и говорим о задаче на экстремум функционала f.

Отметим еще, что если $\widehat{x}$ — решение задачи (P) на минимум (максимум), то ясно, что $\widehat{x}$ — решение аналогичной задачи на максимум (минимум) с функционалом $-f$ вместо f.

Точки из множества ограничений C называются *допустимыми* в задаче (P). Если $C = X$, то задача (P) называется задачей *без ограничений*.

П. Ферма в 1638 году иллюстрировал свой метод нахождения экстремумов на решении следующей задачи: *найти прямоугольный треугольник наибольшей площади при данной сумме длин его катетов.* Если a — сумма длин катетов, а x — длина одного из них, то формализованная постановка выглядит так $x(a-x)/2 \to \max,\ 0 \le x \le a$.

При решении многих конкретных задач нашей целью будет нахождение глобальных экстремумов, но для этого предварительно приходиться исследовать задачу на наличие локальных экстремумов (т. е. локальных минимумов и максимумов). Если в X определено понятие "окрестности точки", то точка $\widehat{x} \in C$ называется *локальным минимумом* (*максимумом*) в задаче (P), если существует такая ее окрестность U, что $f(x) \ge f(\widehat{x})$ $(f(x) \le f(\widehat{x}))$ для всех допустимых $x \in U$ (т. е. для всех $x \in C \cap U$).

Одна из основных целей курса — изложить единый взгляд на необходимые условия экстремума первого порядка для различных классов экстремальных задач (математического и выпуклого программирования, вариационного исчисления и оптимального управления). С этой целью мы рассматриваем более общую ситуацию, связанную с исследованием экстремальных и внутренних точек отображений, и получаем необходимые условия существования экстремальной точки отображения и достаточные условия существования внутренней точки отображения. Оказывается, что необходимые условия экстремума во всех перечисленных классах задач являются непосредственными следствиями этих утверждений.

Глава 1. Аппарат теории экстремальных задач

§ 1.1. Пространство $\mathbb{R}^n$ и двойственное (сопряженное) к нему .

Пространство линейных операторов. Сопряженный оператор. Формула действия сопряженного оператора.

Пусть n — натуральное число. 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 $\mathbb{R}^n$ будем записывать так $x = (x_1, \dots, x_n)^T$, где ^{T} обозначает *транспонирование* (перевод строки в столбец и наоборот). 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 $a = (a_1, \dots, a_n)$ — вектор-строка из n действительных чисел. Для каждого $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ обозначим $\langle a, x \rangle = \sum_{i=1}^n a_i x_i$. Это матричное произведение вектор-строки a на вектор-столбец x, которое иногда называют *внутренним произведением*. Ясно, что отображение $x \mapsto \langle a, x \rangle$ есть линейная функция (или говорят, линейный функционал) на $\mathbb{R}^n$. Легко видеть, что и любой линейный функционал l на $\mathbb{R}^n$ задается подобным образом с $a = (l(e_1), \dots, l(e_n))$, где $e_1 = (1, 0, \dots, 0)^T, \dots, e_n = (0, \dots, 0, 1)^T$ — *стандартный базис* в $\mathbb{R}^n$. Таким образом, если обозначить через

$\mathbb{R}_1^n$ множество, элементы которого суть те же наборы из n действительных чисел, но расположенные в строку (с аналогичными операциями сложения и умножения на числа), то его можно отождествить с пространством всех линейных функционалов на $\mathbb{R}^n$, которое называют *сопряженным* или *двойственным* пространством к $\mathbb{R}^n$ и которое будем обозначать $(\mathbb{R}^n)^*$.

Далее, каждому $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ можно сопоставить линейный функционал $a \mapsto \langle a, x\rangle$ на $(\mathbb{R}^n)^*$ и тогда совершенно аналогично устанавливается, что сопряженное (двойственное) пространство к $(\mathbb{R}^n)^*$ можно отождествить с $\mathbb{R}^n$ (говорят еще, что второе сопряженное к $\mathbb{R}^n$ совпадает с ним самим, т. е. $((\mathbb{R}^n)^*)^* = (\mathbb{R}^n)^{**} = \mathbb{R}^n$).

Пусть $\Lambda\colon \mathbb{R}^n \to \mathbb{R}^m$ — линейный оператор. Мы можем отождествить этот оператор с его матрицей в стандартных базисах $e_1, \dots, e_n$ и $e'_1, \dots, e'_m$ в $\mathbb{R}^n$ и $\mathbb{R}^m$ соответственно, действительно, если $\Lambda e_j = \sum_{i=1}^m a_{ij} e'_i$, $j = 1, \dots, n$, то матрицей оператора Λ однозначно определяет его действие и называется матрица (мы ее обозначаем той же буквой) $\Lambda = (a_{ij})_{1\le i\le m, 1\le j\le n}$ размера $m \times n$. В этом случае Λx — произведение матрицы Λ на вектор $x = (x_1, \dots, x_n)^T$.

Каждому линейному оператору $\Lambda\colon \mathbb{R}^n \to \mathbb{R}^m$ можно сопоставить линейный оператор $\Lambda^*\colon (\mathbb{R}^m)^* \to (\mathbb{R}^n)^*$, который называется *сопряженным оператором* к Λ, и который линейному функционалу $y^* \in (\mathbb{R}^m)^*$ сопоставляет линейный функционал $\Lambda^* y^* \in (\mathbb{R}^n)^*$, действующий по правилу: $\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle$, $x \in \mathbb{R}^n$.

Так как можно считать, что $y^* = (y_1^*, \dots, y_m^*)$ и $\Lambda = (a_{ij}), 1 \le i \le m$, $1 \le j \le n$, то справедлива формула $\Lambda^* y^* = y^* \Lambda$, в которой слева стоит действие сопряженного оператора Λ^* на элемент $y^* \in (R^m)^*$, а справа умножение строки y^* на матрицу Λ. Действительно, если $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, то $\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle = \sum_{i=1}^m y_i^* \left(\sum_{j=1}^n a_{ij} x_j\right) = \sum_{j=1}^n \left(\sum_{i=1}^m y_i^* a_{ij}\right) x_j$. В скобках стоит произведение $y^* A$.

Мы будем встречаться с ситуацией, когда пространство $\mathbb{R}^n$ представлено как произведение пространств: $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $n = n_1 + n_2$, и таким образом, если $x \in \mathbb{R}^n$, то $x = (x_1, x_2)$, где $x_i \in \mathbb{R}^{n_i}$, $i = 1, 2$.

Предложение 1. *Всякий линейный функционал $x^* \in (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})^*$ однозначно представим в виде*

$$\langle x^*, (x_1, x_2)\rangle = \langle x_1^*, x_1\rangle + \langle x_2^*, x_2\rangle.$$

где $x_i^* \in (\mathbb{R}^{n_i})^*$, $i = 1, 2$.

Доказательство. Пусть $x^* \in (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2})^*$. Обозначим через $x_i^* \in (\mathbb{R}^{n_i})^*$, $i = 1, 2$, функционалы, действующие по правилу: $\langle x_1^*, x_1 \rangle = \langle x^*, (x_1, 0) \rangle$ и $\langle x_2^*, x_2 \rangle = \langle x^*, (0, x_2) \rangle$. Отсюда следует нужное представление для x^*. Его единственность проверяется без труда. $\square$

Если $\Lambda \colon \mathbb{R}^n \to \mathbb{R}^m$ — линейный оператор, то его образ и ядро определяются и обозначаются соответственно следующим образом:

$$\operatorname{Im} \Lambda = \{\, y \in \mathbb{R}^m \mid \exists\, x \in \mathbb{R}^n : \Lambda x = y \,\}$$

и

$$\operatorname{Ker} \Lambda = \{\, x \in \mathbb{R}^n : \Lambda x = 0 \,\}.$$

§ 1.2. Открытые и замкнутые множества.

Компактность. Лемма о центрированной системе. Теорема Вейерштрасса.

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. Длина вектора удовлетворяет следующим свойствам: $|x| \geq 0$ и $|x| = 0$ тогда и только тогда, когда $x = 0$, $|\alpha x| = |\alpha||x|$ для любых $x \in \mathbb{R}$ и $\alpha \in \mathbb{R}$, $|x + y| \leq |x| + |y|$ для любых $x, y \in \mathbb{R}$.

Величина $d(x, y) = |x - y|$ называется *расстоянием* между векторами x и y.

Пусть $\widehat{x} \in \mathbb{R}^n$ и $\delta > 0$. Множество $U_{\mathbb{R}^n}(\widehat{x}, \delta) = \{\, x \in \mathbb{R}^n \mid |x - \widehat{x}| < \delta \,\}$ называется *открытым* шаром с центром в точке $\widehat{x}$ радиуса δ.

Пусть $A \subset \mathbb{R}^n$ и $x \in A$. Говорят, что x — *внутренняя точка* A, если x входит в A вместе с некоторым открытым шаром с центром в этой точке. Множество внутренних точек A называется *внутренностью* A и обозначается $\operatorname{int} \mathrm{A}$. Множество $G \subset \mathbb{R}^n$ называется *открытым*, если оно совпадает со свое внутренностью.

Упражнение 1.2. *Доказать, что открытый шар — открытое множество.*

Упражнение 1.3. *Доказать, что объединение любого и пересечение конечного числа открытых множеств — открытое множество.*

Окрестностью множества называется любое открытое множество, содержащее данное множество. В частности, *окрестностью точки* называется любое открытое множество, содержащее данную точку.

Точка $\widehat{x} \in \mathbb{R}^n$ называется *граничной точкой множества* A, если в любой ее окрестности содержатся точки из A и из $\mathbb{R}^n \setminus A$. Множество всех граничных точек множества обозначаем ∂A.

Упражнение 1.4. *Доказать, что если $x \in A$, то либо $x \in \operatorname{int} A$, либо $x \in \partial A$.*

Множество $F \subset \mathbb{R}^n$ называется *замкнутым*, если $\mathbb{R}^n \setminus F$ — открытое множество.

Упражнение 1.5. *Доказать, что шар (иногда говорят замкнутый шар) $B_{\mathbb{R}^n}(\widehat{x}, \delta) = \{\, x \in \mathbb{R}^n \mid |x - \widehat{x}| \le \delta \,\}$ с центром в точке $\widehat{x}$ радиуса δ — замкнутое множество.*

Упражнение 1.6. *Доказать, что пересечение любого числа и объединение конечного числа замкнутых множеств — замкнутое множество.*

Упражнение 1.7. *Доказать, что подпространства в $\mathbb{R}^n$ — замкнутые множества.*

Пусть $A \subset \mathbb{R}^n$. Точка $x \in \mathbb{R}^n$ называется *предельной точкой множества* A, если каждая ее окрестность имеет непустое пересечение с A. Совокупность всех предельных точек множества A называется *замыканием* A и обозначается $\operatorname{cl} A$. Ясно, что $A \subset \operatorname{cl} A$.

Упражнение 1.8. *Доказать, что* $\operatorname{cl} A$ *— замкнутое множество и что множество A замкнуто тогда и только тогда, когда оно совпадает со своим замыканием.*

Множество $A \subset \mathbb{R}^n$ называется *компактным* (или *компактом*), если из любого его покрытия открытыми множествами можно выделить конечное подпокрытие.

Семейство множеств называется *центрированной системой*, если любое его конечное подсемейство имеет непустое пересечение.

Лемма (о центрированной системе)**.** *Множество компактно тогда и только тогда, когда любая его центрированная система замкнутых подмножеств имеет непустое пересечение.*

Доказательство. Пусть множество A компактно и A_i, $i \in J$, — центрированная система его замкнутых подмножеств. Предположим, что $\cap_{i \in J} A_i = \emptyset$. Тогда, так как $A = A \setminus \cap_{i \in J} A_i = \cup_{i \in J}(A \setminus A_i) \subset \cup_{i \in J}(\mathbb{R}^n \setminus A_i)$, то открытые множества $\mathbb{R}^n \setminus A_i$, $i \in J$, образуют покрытие A. Поскольку A компактно, то существует конечное подсемейство $\mathbb{R}^n \setminus A_{i_1}, \ldots, \mathbb{R}^n \setminus A_{i_m}$, которое покрывает A. С другой стороны, по условию, $\cap_{k=1}^m A_{i_k} \ne \emptyset$ и мы имеем $\cap_{k=1}^m A_{i_k} \subset A \subset$

$\cup_{k=1}^{m}(\mathbb{R}^n \setminus A_{i_k}) = \mathbb{R}^n \setminus \cap_{k=1}^{m} A_{i_k}$. Противоречие доказывает утверждение.

Доказательство в другую сторону аналогично. $\square$

Упражнение 1.9. *Доказать, что множество $A \subset \mathbb{R}^n$ компактно тогда и только тогда, когда из любой последовательности его элементов можно выделить подпоследовательность, сходящуюся к элементу из A.*

Упражнение 1.10. *Доказать, что множество в $\mathbb{R}^n$ компактно тогда и только тогда, когда оно ограничено (т. е. содержится в некотором шаре) и замкнуто.*

Упражнение 1.11. *Доказать, что если K — компакт, а F — замкнутое множество, то их алгебраическая сумма $K+F = \{\, z = x + y \mid x \in K,\ y \in F \,\}$ — замкнутое множество.*

Выше мы говорили о свойствах пространства $\mathbb{R}^n$ (норма элемента, открытое множество, замкнутое множество, компактность). Все это дословно переносится на пространство $(\mathbb{R}^n)^*$.

Напомним понятие непрерывного отображения. Пусть A — подмножество $\mathbb{R}^n$ и задано отображение $F\colon A \to \mathbb{R}^m$. Это равносильно тому, что на A задано m функций. Действительно, координаты $F(x)$ суть m функций на A. С другой стороны, если на A заданы функции $f_1, \dots, f_m$, то они определяют отображение $F\colon A \to \mathbb{R}^m$ по правилу $F(x) = (f_1(x), \dots, f_m(x))^T$, $x \in A$. В этом смысле многие вопросы, связанные с отображениями можно свести к соответствующим вопросам для функций. Однако часто удобно (и полезно, имея в ввиду обобщения на бесконечномерный случай) работать именно с отображениями.

Отображение $F\colon A \to \mathbb{R}^m$ называется *непрерывным в точке* $\widehat{x} \in A$, если для любого $\varepsilon > 0$ найдется такое $\delta = \delta(\varepsilon) > 0$, что $|F(x) - F(\widehat{x})| < \varepsilon$ для всех $x \in A \cap U_{\mathbb{R}^n}(\widehat{x}, \delta)$, или равносильно: для любой последовательности $\{x_k\}$ точек из A, сходящейся к $\widehat{x}$ последовательность $\{F(x_k)\}$ сходится к $F(\widehat{x})$.

Говорят, что отображение F *непрерывно на* A, если оно непрерывно в каждой точке A.

Если $m = 1$, то мы говорим, естественно, о непрерывной функции.

Упражнение 1.12. *Доказать, что отображение $F\colon \mathbb{R}^n \to \mathbb{R}^m$ непрерывно тогда и только тогда, когда прообраз любого открытого множества открыт.*

Упражнение 1.13. *Доказать, что функции* $x \to |x|$ *и* $x \to \langle x^*, x\rangle + \alpha$, *где* $x^* \in (\mathbb{R}^n)^*$ *и* $\alpha \in \mathbb{R}$, *непрерывны на* $\mathbb{R}^n$.

Теорема (Вейерштрасса)**.** *Функция непрерывная на компакте в* $\mathbb{R}^n$ *достигает на нем своего максимального и минимального значений.*

Доказательство. Пусть A — компактное подмножество $\mathbb{R}^n$ и $f : A \to \mathbb{R}$ — непрерывная функция. Обозначим $\gamma = \sup\{\, f(x) \mid x \in A\,\}$. По определению верхней грани (конечной или бесконечной) существует последовательность $\{x_k\}$ элементов из A такая, что $\{f(x_k)\}$ сходится к γ. Из $\{x_k\}$ можно выделить подпоследовательность $\{x_{k_l}\}$, которая сходится к $\overline{x} \in A$. Тогда последовательность $\{f(x_{k_l})\}$ сходится к $f(\overline{x}) = \gamma$, т. е. γ конечно и тем самым $\overline{x}$ — точка, где f принимает свое максимальное значение. Аналогичные рассуждения относительно минимума функции. $\square$

В предыдущем пункте было определено понятие линейного оператора из $\mathbb{R}^n$ в $\mathbb{R}^m$.

Упражнение 1.14. *Доказать, что линейный оператор* Λ *из* $\mathbb{R}^n$ *в* $\mathbb{R}^m$ *есть непрерывное отображение.*

Обозначим через $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ множество всех таких операторов. Это векторное пространство с обычными (поэлементными) операциями сложения матриц и умножения их на числа. Каждому $\Lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ сопоставим число $\|\Lambda\| = \sup_{|x|\le 1} |\Lambda x|$[2], которое называется нормой оператора Λ. Простая проверка показывает, что выполняются следующие условия: (a) $\|\Lambda\| \ge 0$ для любого $\Lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ и $\|\Lambda\| = 0$ тогда и только тогда, когда $\Lambda = 0$; (b) $\|\alpha\Lambda\| = |\alpha|\|\Lambda\|$ для любых $\alpha \in \mathbb{R}$ и $\Lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$; (c) $\|\Lambda_1 + \Lambda_2\| \le \|\Lambda_1\| + \|\Lambda_2\|$ для любых $\Lambda_1, \Lambda_2 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Следовательно, введенная таким образом функция $\Lambda \mapsto \|\Lambda\|$ является (операторной) нормой в $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Ясно также, что $|\Lambda x| \le \|\Lambda\||x|$ для любого $x \in \mathbb{R}^n$.

Данное определение нормы в $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ позволяет (также как и в $\mathbb{R}^n$) ввести в этом пространстве понятия открытого и замкнутого множества, окрестности точки и т. д.

§ 1.3. Выпуклые множества и конусы. Теоремы отделимости.

[2]На самом деле здесь можно поставить max, так как по теореме Вейерштрасса непрерывная функция $x \to |\Lambda x|$ достигает своего максимума на шаре $B_{\mathbb{R}^n}(0, 1)$.

Нормальный конус. Критерий граничной точки выпуклого множества.

Пусть $x, y \in \mathbb{R}^n$. Множество $[x,y] = \{z \in \mathbb{R}^n : z = (1-\alpha)x + \alpha y,\ 0 \le \alpha \le 1\}$ называется *отрезком* (соединяющим точки x и y).

Непустое множество $A \subset \mathbb{R}^n$ называется *выпуклым,* если вместе с любыми двумя своими точками x и y оно содержит отрезок $[x,y]$.

Пустое множество выпукло по определению.

Следующие свойства выпуклых множеств непосредственно следуют из определений (проверьте).

Если $\{A_i\}_{i\in\mathcal{J}}$ — произвольное семейство выпуклых подмножеств $\mathbb{R}^n$, то $\bigcap_{i\in\mathcal{J}} A_i$ — выпуклое множество.

Если $\{A_i\}_{i=1}^n$ — конечное семейство выпуклых подмножеств $\mathbb{R}^n$, то $A_1 + \dots + A_n := \{x \in \mathbb{R}^n \mid x = x_1 + \dots + x_n,\ x_i \in A_i,\ i = 1, \dots, n\}$ — выпуклое множество.

Если A — выпуклое подмножество $\mathbb{R}^n$ и $\lambda \in \mathbb{R}$, то множество $\lambda A := \{\lambda x \mid x \in A\}$ выпукло.

Упражнение 1.15. *Доказать, что открытые и замкнутые шары в $\mathbb{R}^n$ — выпуклые множества.*

Пусть A — непустое подмножество $\mathbb{R}^n$. Наименьшее выпуклое множество, содержащее A (т. е. пересечение всех выпуклых множеств, содержащих A) называется *выпуклой оболочкой множества* A и обозначается $\operatorname{co} \mathrm{A}$.

Упражнение 1.16. *Доказать, что* $\operatorname{co} \mathrm{A}$ *состоит из выпуклых комбинаций элементов из A, т. е. векторов вида $\alpha_1 x_1 + \ldots + \alpha_m x_m$, где $x_i \in A$, $\alpha_i \ge 0$, $1 \le i \le m$, и $\sum_{i=1}^m \alpha_i = 1$.*

Пусть $\lambda \in (\mathbb{R}^n)^*$, $\lambda \ne 0$ и $\gamma \in \mathbb{R}$. Множество $H = H(\lambda, \gamma) = \{\, x \in \mathbb{R}^n : \langle \lambda, x \rangle = \gamma \,\}$ называется *гиперплоскостью.* Гиперплоскость порождает два полупространства $H_+(\lambda,\gamma) = \{\, x \in \mathbb{R}^n : \langle \lambda, x \rangle \le \gamma \,\}$ и $H_-(\lambda,\gamma) = \{\, x \in \mathbb{R}^n : \langle \lambda, x \rangle \ge \gamma \,\}$.

Упражнение 1.17. *Доказать, что гиперплоскости и полупространства — выпуклые замкнутые множества.*

Непустое множество K в $\mathbb{R}^n$ называется *конусом*, если из включения $x \in K$ следует, что $\alpha x \in K$ для любого $\alpha > 0$.

Упражнение 1.18. *Доказать, что конус C является выпуклым тогда и только тогда, когда $K + K \subset K$.*

Пусть A — непустое подмножество $\mathbb{R}^n$ и $\widehat{x} \in A$. Множество $N_A(\widehat{x}) = \{\, \lambda \in (\mathbb{R}^n)^* : \langle \lambda, x - \widehat{x} \rangle \le 0,\ \forall\, x \in A \,\}$ называется *нормальным конусом ко множеству A в точке $\widehat{x}$.*

Упражнение 1.19. *Доказать, что $N_A(\widehat{x})$ — выпуклый замкнутый конус.*

Пусть A и B — непустые подмножества $\mathbb{R}^n$. Говорят, что эти множества *отделимы*, если существует такая гиперплоскость, что A и B принадлежат различным полупространствам, порожденным этой гиперплоскостью.

Данное геометрическое определение отделимости равносильно, очевидно, следующему алгебраическому: множества A и B отделимы, если существует такой ненулевой элемент $\lambda \in (\mathbb{R}^n)^*$, что

$$\sup_{a\in A} \langle \lambda, a\rangle \leq \inf_{b\in B} \langle \lambda, b\rangle.$$

Если неравенство строгое, то говорят, что множества A и B *строго отделимы.*

Теорема (Первая теорема отделимости)**.** *Пусть A и B — непустые выпуклые подмножества $\mathbb{R}^n$ и $A \cap B = \emptyset$. Тогда A и B отделимы.*

Эту теорему здесь не доказываем.

Теорема (Вторая теорема отделимости)**.** *Пусть A — непустое выпуклое замкнутое подмножество $\mathbb{R}^n$ и $b \notin A$. Тогда множество A и точка b строго отделимы.*

Доказательство. Так как A замкнуто, то дополнение открыто и поэтому существует такое $r > 0$, что $A\cap U_{\mathbb{R}^n}(b, r) = \emptyset$. Согласно первой теореме отделимости найдется ненулевое $\lambda \in (\mathbb{R}^n)^*$ такое, что $\sup_{x\in A}\langle \lambda, x\rangle \leq \inf_{x\in U_{\mathbb{R}^n}(b,r)}\langle \lambda, x\rangle$. Но $\inf_{x\in U_{\mathbb{R}^n}(b,r)}\langle \lambda, x\rangle < \langle \lambda, b\rangle$, так как ненулевой линейный функционал не может достигать минимума во внутренней точке (проверьте это). Следовательно множество A и точка b строго отделимы. □

Ниже, используя первую теорему отделимости, мы докажем одно утверждение, которое нам понадобится в дальнейшем.

Предложение 2. *Пусть A — выпуклое подмножество $\mathbb{R}^n$ и $\widehat{x} \in A$. Тогда $\widehat{x} \in \partial A$ в том и только в том случае, если $N_A(\widehat{x}) \neq \{0\}$.*

Доказательство. Необходимость. Пусть $\widehat{x} \in \partial A$. Найдется последовательность $x_k \in \mathbb{R}^n \setminus A$, $k \in \mathbb{N}$, такая, что $\lim_{k\to\infty} x_k = \widehat{x}$. Согласно первой теореме отделимости существует последовательность ненулевых векторов $\lambda_k \in (\mathbb{R}^n)^*$, $k \in \mathbb{N}$, обладающая тем свойством, что

$$\langle \lambda_k, x\rangle \leq \langle \lambda_k, x_k\rangle, \qquad \forall\, x \in A. \tag{1}$$

Можно считать, что $|\lambda_k| = 1$, $k \in \mathbb{N}$, и что последовательность λ_k сходится к некоторому λ. Ясно, что $|\lambda| = 1$. Переходя к пределу в (1) при $k \to \infty$, получаем, что $\langle \lambda, x - \widehat{x} \rangle \le 0$ для всех $x \in A$, т. е. $\lambda \in N_A(\widehat{x})$.

Достаточность. Пусть $\lambda \in N_A(\widehat{x})$ и $\lambda \ne 0$. Тогда существует такой вектор $y \in \mathbb{R}^n$, что $\langle \lambda, y \rangle > 0$ и значит, $\alpha y \notin A - \widehat{x}$, или равносильно, $\alpha y + \widehat{x} \notin A$ для любого $\alpha > 0$, т. е. $\widehat{x} \in \partial A$. $\square$

§ 1.4. Дифференциальное исчисление.

Дифференцируемость, непрерывная дифференцируемость, строгая дифференцируемость. Частные производные. Вторая производная.

Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$. Говорят, что функция $f : U \to \mathbb{R}$ *дифференцируема в точке* $\widehat{x}$, если существует линейный функционал на $\mathbb{R}^n$, т. е. вектор $a = (a_1, \ldots, a_n) \in (\mathbb{R}^n)^*$ такой, что для всех $h \in \mathbb{R}^n$, для которых $\widehat{x} + h \in U$ справедливо представление

$$f(\widehat{x} + h) = f(\widehat{x}) + \langle a, h \rangle + r(h),$$

где $|r(h)|/|h| \to 0$ при $h \to 0$, т. е. $r(h) = o(h)$. Вектор a называется *производной функции* f *в точке* $\widehat{x}$ и обозначается $f'(\widehat{x})$.

Упражнение 1.20. *Доказать, что производная определена однозначно.*

Из данного определения легко следует (беря в качестве h векторы $(h_1, 0, \ldots, 0)^T, \ldots, (0, \ldots, 0, h_n)^T$), что a_i есть частная производная функции f по x_i в точке $\widehat{x}$, $1 \le i \le n$. Таким образом, $f'(\widehat{x}) = (\partial f(\widehat{x})/\partial x_1, \ldots, \partial f(\widehat{x})/\partial x_n)$.

Упражнение 1.21. *Доказать, что функция, дифференцируемая в данной точке, непрерывна в этой точке.*

Существование частных производных функции в данной точке не гарантирует ее дифференцируемость и даже непрерывность в данной точке.

Упражнение 1.22. *Постройте функцию, у которой существуют частные производные в точке, но сама функция в этой точке разрывна.*

Упражнение 1.23. *Пусть A — симметричная матрица размера $n \times n$, $b \in (\mathbb{R}^n)^*$, $c \in \mathbb{R}$ и функция $f \colon \mathbb{R}^n \to \mathbb{R}$ определена по правилу: $f(x) = x^T A x + \langle b, x \rangle + c$. Найдите ее производную для каждого $x \in \mathbb{R}^n$.*

Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$. Говорят, что отображение $F\colon U \to \mathbb{R}^m$ *дифференцируемо в точке* $\widehat{x}$, если существует линейный оператор из $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, т. е. матрица Λ размера $m \times n$, такая, что для всех $h \in \mathbb{R}^n$, для которых $\widehat{x} + h \in U$, справедливо представление

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

где $|\rho(h)|/|h| \to 0$ при $h \to 0$, т. е. $\rho(h) = o(h)$. Матрица Λ называется *производной отображения* F *в точке* $\widehat{x}$ и обозначается $F'(\widehat{x})$.

Упражнение 1.23. *Доказать, что производная определена однозначно.*

Упражнение 1.24. *Пусть* $\widehat{x} \in \mathbb{R}^n$. *Чему равна производная отображения* $\Lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ *в данной точке?*

Если на U определены функции f_i, $i = 1, \ldots, m$, и отображение $F\colon U \to \mathbb{R}^m$ определено по правилу: $F(x) = (f_1(x), \ldots, f_m(x))^T$, то легко проверить, что отображение F дифференцируемо в точке $\widehat{x}$ тогда и только тогда, когда функции f_i, $1 \le i \le m$, дифференцируемы в $\widehat{x}$ (проверьте это). При этом строки матрицы $F'(\widehat{x})$ суть производные $f_1'(\widehat{x}), \ldots, f_k'(\widehat{x})$ и тем самым $F'(\widehat{x}) = (\partial f_i(\widehat{x})/\partial x_j)_{1 \le i \le m,\, 1 \le j \le n}$. Эту матрицу называют *матрицей Якоби* отображения F в точке $\widehat{x}$.

Если отображение F дифференцируемо в каждой точке U, то говорят, что F *дифференцируемо на* U.

Пусть отображение $F\colon U \to \mathbb{R}$ дифференцируемо на U. Тогда определено отображение $F'\colon U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, называемое *производной отображения* F. Если производная F непрерывна в точке $\widehat{x}$ (на U), то говорят, что отображение F *непрерывно дифференцируемо в* $\widehat{x}$ (на U).

Упражнение 1.25*. *Доказать, что если отображение* F *задается набором функций* $f_1, \ldots, f_m$, *то непрерывная дифференцируемость* F *на* U *равносильно тому, что все частные производные этих функций непрерывны на* U.

Говорят, что отображение $F\colon U \to \mathbb{R}^m$ *строго дифференцируемо в точке* $\widehat{x}$, если существует такой линейный оператор из $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, т. е. матрица Λ размера $m \times n$, что для любого $\varepsilon > 0$ найдется $\delta = \delta(\varepsilon) > 0$, обладающее тем свойством, что для всех $x, x' \in U_{\mathbb{R}^n}(\widehat{x}, \delta)$ выполняется неравенство

$$|F(x) - F(x') - \Lambda(x - x')| \le \varepsilon |x - x'|.$$

Если $x' = \widehat{x}$, то получаем определение, равносильное дифференцируемости F в точке $\widehat{x}$ и значит, $\Lambda = F'(\widehat{x})$. Таким образом, строго дифференцируемое отображение в $\widehat{x}$ дифференцируемо в этой точке.

Если $m = 1$, т. е. F — функция, то получаем определение строгой дифференцируемости функции в точке.

Если отображение F порождено набором функций, то простая проверка показывает, что строгая дифференцируемость F в точке $\widehat{x}$ равносильна строгой дифференцируемости в $\widehat{x}$ каждой из этих функций.

Упражнение 1.26. *Доказать, что если отображение строго дифференцируемо в точке $\widehat{x}$, то для каждого $\varepsilon > 0$ оно непрерывно в соответствующей окрестности $U_{\mathbb{R}^n}(\widehat{x}, \delta)$.*

Упражнение 1.27. *Построить пример функции, которая дифференцируема в данной точке, но не строго дифференцируема в ней.*

Отметим, что если отображение непрерывно дифференцируемо в точке, то оно строго дифференцируемо в этой точке. Это простое следствие теоремы о среднем, но более подробно на этом останавливаться не будем.

Теорема (о производной суперпозиции отображений)**.** *Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$, отображение $F_1\colon U \to \mathbb{R}^m$ дифференцируемо (непрерывно дифференцируемо, строго дифференцируемо) в $\widehat{x}$, V — окрестность точки $F_1(\widehat{x})$ и отображение $F_2\colon V \to \mathbb{R}^k$ дифференцируемо (непрерывно дифференцируемо, строго дифференцируемо) в $\widehat{y} = F_1(\widehat{x})$. Тогда отображение $F = F_2 \circ F_1\colon U \to \mathbb{R}^k$ дифференцируемо (непрерывно дифференцируемо, строго дифференцируемо) в точке $\widehat{x}$ и $F'(\widehat{x}) = F_2'(\widehat{y})F_1'(\widehat{x})$.*

При выводе необходимых условий экстремума будут встречаться ситуации, когда отображение определено не на всей окрестности точки. Естественным образом определяется производная отображения и в этом случае (простейший пример — производные функции одного переменного слева и справа в данной точке).

Пусть K — выпуклое подмножество $\mathbb{R}^n$ и $\widehat{x} \in K$. Скажем, что отображение $F\colon K \to \mathbb{R}^m$ *дифференцируемо в точке $\widehat{x}$ относительно K*, если найдется такой оператор из $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, т. е. матрица Λ размера $m \times n$, что для всех $h \in \mathbb{R}^n$, для которых $\widehat{x} + h \in K$ справедливо представление

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

где $|\rho(h)|/|h| \to 0$ при $h \to 0$.

Оператор Λ будем называть производной отображения F в точке $\widehat{x}$ относительно множества K и обозначать $F'_K(\widehat{x})$ или просто $F'(\widehat{x})$, если ясно относительно какого множества берется производная.

Заметим, что если $\operatorname{int} K \neq \emptyset$, то производная $F'_K(\widehat{x})$ определена однозначно. Действительно, пусть сначала $\widehat{x} \notin \operatorname{int} K$. Возьмем произвольное $x_0 \in \operatorname{int} K$. Тогда существует такое $r > 0$, что $x_0 + U_{\mathbb{R}^n}(0,r) \subset K$. Если $y \in \mathbb{R}^n$ и $|y| = 1$, то $(r/2)y \in U_{\mathbb{R}^n}(0,r)$ и поэтому $x_0 + (r/2)y \in K$. Так как K выпукло, то $(1-\alpha)\widehat{x} + \alpha(x_0 + (r/2)y) = \widehat{x} + \alpha(x_0 - \widehat{x} + (r/2)y) \in K$ для всех $0 < \alpha < 1$.

Если бы было два различных представления Λ_1 и Λ_2 для производной, то мы получили бы в пределе при $\alpha \to 0$, что $(\Lambda_1 - \Lambda_2)y = (2/r)(\Lambda_1 - \Lambda_2)(\widehat{x} - x_0)$, т. е. оператор $\Lambda_1 - \Lambda_2$ на единичной сфере постоянен и поэтому это нулевой оператор.

Если $\widehat{x} \in \operatorname{int}$ K, то это обычная производная (определенная выше) отображения F в точке $\widehat{x}$.

Строгая дифференцируемость отображения $F\colon K \to \mathbb{R}^m$ относительно выпуклого множества K определяется аналогично: для любого $\varepsilon > 0$ неравенство в определении строгой дифференцируемости должно выполнять для всех $x, x' \in U_{\mathbb{R}^n}(\widehat{x}, \delta) \cap K$.

Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$ и $F_i\colon U \to \mathbb{R}^{n_i}$, $i = 1, 2$. Рассмотрим отображение $F = (F_1, F_2)\colon U \to \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, действующее по правилу: $F(x) = (F_1(x), F_2(x))$.

Предложение 3. *Отображение F дифференцируемо в точке $\widehat{x}$ тогда и только тогда, когда отображения F_1 и F_2 дифференцируемы в точке $\widehat{x}$ и при этом $F'(\widehat{x}) = (F'_1(\widehat{x}), F'_2(\widehat{x}))$.*

Доказательство. Пусть отображения F_1 и F_2 дифференцируемы в точке $\widehat{x}$. Тогда по определению

$$F(\widehat{x}+h) = (F_1(\widehat{x}+h), F_2(\widehat{x}+h)) = (F_1(\widehat{x}), F_2(\widehat{x})) + (F'_1(\widehat{x}), F'_2(\widehat{x}))h \\ + (r_1(h), r_2(h)),$$

где $|r_i(h)|/|h| \to 0$ при $h \to 0$, $i = 1, 2$. Следовательно, отображение F дифференцируемо в точке $\widehat{x}$ и $F'(\widehat{x}) = (F'_1(\widehat{x}), F'_2(\widehat{x}))$.

В другую сторону рассуждения аналогичны (проделайте это). $\square$

Пусть W — окрестность точки $(\widehat{x}, \widehat{y}) \in \mathbb{R}^n \times \mathbb{R}^k$ и $F\colon W \to \mathbb{R}^m$. Если отображение $x \mapsto F(x, \widehat{y})$ (определенное на проекции W на $\mathbb{R}^n$) дифференцируемо в точке $\widehat{x}$, то соответствующую производную называют *частной производной по x отображения F в точке*

$(\widehat{x}, \widehat{y})$ и обозначают $F_x(\widehat{x}, \widehat{y})$. Аналогично частную производную по y отображения F в точке $(\widehat{x}, \widehat{y})$ обозначают $F_y(\widehat{x}, \widehat{y})$.

Предложение 4. *Если отображение $F\colon W \to \mathbb{R}^m$ дифференцируемо в точке $(\widehat{x}, \widehat{y})$, то оно имеет в этой точке частные производные по x и y, и для всех $\xi \in \mathbb{R}^n$ и $\eta \in \mathbb{R}^k$ справедливо представление*

$$F'(\widehat{x}, \widehat{y})[\xi, \eta] = F_x(\widehat{x}, \widehat{y})[\xi] + F_y(\widehat{x}, \widehat{y})[\eta]^3$$

Доказательство. Так как F дифференцируемо в точке $(\widehat{x}, \widehat{y})$, то, в частности, справедливо представление $F(\widehat{x} + h, \widehat{y}) = F(\widehat{x}, \widehat{y}) + F'(\widehat{x}, \widehat{y})(h, 0) + r(h)$, где $|r(h)|/|h| \to 0$ при $h \to 0$. Отсюда, в силу единственности, следует, что линейный оператор $\xi \mapsto F'(\widehat{x}, \widehat{y})[\xi, 0]$ из $\mathbb{R}^n$ в $\mathbb{R}^m$ совпадает с $F_x(\widehat{x}, \widehat{y})$. Аналогично, линейный оператор $\eta \mapsto F'(\widehat{x}, \widehat{y})[0, \eta]$ из $\mathbb{R}^k$ в $\mathbb{R}^m$ совпадает с $F_y(\widehat{x}, \widehat{y})$. Но тогда $F'(\widehat{x}, \widehat{y})[\xi, \eta] = F'(\widehat{x}, \widehat{y})([\xi, 0] + [0, \eta]) = F_x(\widehat{x}, \widehat{y})[\xi] + F_y(\widehat{x}, \widehat{y})[\eta]$. $\square$

Пусть U — окрестность точки $\widehat{x}$ и отображение $F\colon U \to \mathbb{R}^m$ дифференцируемо на U. Если отображение $F'\colon U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ дифференцируемо в $\widehat{x}$, то говорят, что F дважды дифференцируемо в точке $\widehat{x}$ и соответствующую (вторую) производную обозначают $F''(\widehat{x})$. Ясно, что $F''(\widehat{x}) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$. Пусть $h_1 \in \mathbb{R}^n$. Тогда $F''(\widehat{x})h_1 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Для любого $h_2 \in \mathbb{R}^n$ вектор $F''(\widehat{x})h_1[h_2]$ принадлежит $\mathbb{R}^m$. Ясно, что отображение из $\mathbb{R}^n \times \mathbb{R}^n$ в $\mathbb{R}^m$, действующее по правилу $(h_1, h_2) \mapsto F''(\widehat{x})h_1[h_2]$, есть билинейное отображение. Таким образом, $F''(\widehat{x})$ можно рассматривать как билинейное отображение, действие которого на элементе (h_1, h_2) будем записывать так $F''(\widehat{x})[h_1, h_2]$.

Упражнение 1.28. *Доказать, что $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ изометрически изоморфно пространству всех билинейных отображений $B\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ с нормой $\|B\| = \sup_{|h_1| \le 1, \, |h_2| \le 1} |B[h_1, h_2]|$.*

Пусть отображение F задается функциями $f_i\colon U \to \mathbb{R}$, $i = 1, \ldots, m$. Покажем, что если эти функции имеют на U частные производные первого и второго порядков, непрерывные в $\widehat{x}$, то

$$F''(\widehat{x})[h_1, h_2] = (h_1^T H_1 h_2, \ldots, h_1^T H_m h_2)^T,$$

где $H_k = (\partial^2 f_k(\widehat{x})/\partial x_i \partial x_j)$, $i, j = 1, \ldots, n$, $k = 1, \ldots, m$, — гессианы функций $f_1, \ldots, f_m$ в точке $\widehat{x}$.

Пусть $(h_1, h_2) \in \mathbb{R}^n \times \mathbb{R}^n$. В каждой точке $x \in U$, из доказанного выше, следует, что $F'(x)h_1 = (\langle f_1'(x), h_1 \rangle, \ldots, \langle f_m'(x), h_1 \rangle)^T$. Снова,

---

[3] $F'(\widehat{x}, \widehat{y})[\xi, \eta]$ обозначает действие (значение) оператора $F'(\widehat{x}, \widehat{y})$ на элементе (ξ, η). Аналогично для частных производных.

используя доказанное выше и учитывая симметричность гессианов (следующую из непрерывности частных производных второго порядка), нетрудно проверить, что производная отображения $x \mapsto F'(x)[h_1]$ из U в $\mathbb{R}^n$ в точке $\widehat{x}$ на элементе h_2 имеет вид $(h_1^T H_1 h_2, \dots, h_1^T H_m h_2)^T$. Обозначая это выражение через $B[h_1, h_2]$ (это, очевидно, симметричная билинейная форма), в частности, имеем: $F'(\widehat{x} + \alpha h_2)[h_1] = F'(\widehat{x})[h_1] + \alpha B[h_1, h_2] + o(\alpha)$.

Из определения $F''(\widehat{x})$ следует, что $F'(\widehat{x} + \alpha h_2) = F'(\widehat{x}) + \alpha F''(\widehat{x})h_2 + o(\alpha)$. Применим это к вектору h_1, получим, что $F'(\widehat{x} + \alpha h_2)[h_1] = F'(\widehat{x})[h_1] + \alpha F''(\widehat{x})h_2[h_1] + o(\alpha)[h_1]$. Отсюда и аналогичного соотношения выше сразу вытекает, что $F''(\widehat{x})[h_1, h_2] = B[h_1, h_2]$.

Теорема (Формула Тейлора)**.** *Пусть U — окрестность точки $x \in \mathbb{R}^n$. Если отображение $F\colon U \to \mathbb{R}^m$ имеет вторую производную в точке $\widehat{x}$, то 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 $|r(h)|/|h|^2 \to 0$ при $h \to 0$.

§ 1.5. Метрическая регулярность.

Теорема о возмущении. Метрическая регулярность линейного отображения.

Определение 1. *Пусть M — подмножество $\mathbb{R}^n$. Будем говорить, что отображение $F\colon M \to \mathbb{R}^m$ метрически регулярно в окрестности точки $\widehat{x} \in M$, если найдутся окрестности V_1 и V_2 соответственно точек $\widehat{x}$ и $F(\widehat{x})$ и константа $c > 0$ такие, что*

$$\operatorname{dist}(x, F^{-1}(y)) \le c|y - F(x)| \tag{2}$$

для всех $(x, y) \in (V_1 \cap M) \times V_2$.

В случае, если важна величина константы c, то будем говорить о метрической регулярности с константой c.

Из определения сразу следует, что для любых $(x, y) \in (V_1 \cap M) \times V_2$ и $\varepsilon > 0$ существует $x_\varepsilon(x, y) \in M$ такое, что $F(x_\varepsilon(x, y)) = y$ и при этом

$$|x - x_\varepsilon(x, y)| \le (c + \varepsilon)|y - F(x)|.$$

В частности, если $x = \widehat{x}$, то это есть утверждение о существовании обратной функции у отображения F.

Теорема (о возмущении)**.** *Пусть M — замкнутое подмножество $\mathbb{R}^n$, Φ и G — непрерывные отображения из M в $\mathbb{R}^m$, причем Φ*

метрически регулярно в окрестности точки $\widehat{x} \in M$ *с константой* c*, и существует такая окрестность* V_0 *точки* $\widehat{x}$*, что* G *липшицево на* $V_0 \cap G$ *с константой Липшица* $l < c^{-1}$*. Тогда отображение* $F = \Phi + G$ *метрически регулярно в окрестности точки* $\widehat{x}$*.*

Доказательство. Пусть V_1 и V_2 — окрестности соответственно точек $\widehat{x}$ и $\Phi(\widehat{x})$ из определения метрической регулярности Φ в окрестности $\widehat{x}$, причем можно считать, что $V_1 \subset V_0$. Выберем $\delta_1 > 0$, $\delta_2 > 0$ и $\varepsilon_0 > 0$ так, чтобы $U_{\mathbb{R}^n}(\widehat{x}, \delta_1) \subset V_1$, $U_{\mathbb{R}^m}(\Phi(\widehat{x}), l\delta_1 + \delta_2) \subset V_2$, $\alpha(\varepsilon_0) = (1+\varepsilon_0)lc < 1$ и

$$\frac{(1+\varepsilon_0)c}{1-\alpha(\varepsilon_0)}\,\delta_2 < \frac{\delta_1}{3}\,. \tag{3}$$

В силу непрерывности отображений Φ и G найдется такая окрестность $V_1' \subset U_{\mathbb{R}^n}(\widehat{x}, \delta_1/3)$ точки $\widehat{x}$, что

$$|F(x) - F(\widehat{x})| < \frac{1-\alpha(\varepsilon_0)}{3(1+\varepsilon_0)c}\,\delta_1 \tag{4}$$

для всех $x \in V_1'$.

Пусть $0 < \varepsilon \le \varepsilon_0$ (заметим, что соотношения (3) и (4) тем более справедливы для таких ε). Положим $V_2' = U_{\mathbb{R}^m}(F(\widehat{x}), \delta_2)$ и пусть $(x, y) \in (V_1' \cap M) \times V_2'$. Рассмотрим последовательность x_k, $k \in \mathbb{N}$, заданную рекуррентными соотношениями

$$x_k \in \Phi^{-1}(y - G(x_{k-1}), \quad x_0 = x, \tag{5}$$

и

$$|x_k - x_{k-1}| \le (1+\varepsilon)\mathrm{dist}\,(x_{k-1}, \Phi^{-1}(y - G(x_{k-1})). \tag{6}$$

Покажем, что эта последовательность принадлежит $U_{\mathbb{R}^n}(\widehat{x}, \delta_1) \cap M$ и фундаментальна.

Первое доказываем по индукции. Ясно, что $x_0 = x \in U_{\mathbb{R}^n}(\widehat{x}, \delta_1) \cap M$. Пусть $x_s \in U_{\mathbb{R}^n}(\widehat{x}, \delta_1) \cap M$, $s = 1, \ldots, k$. Покажем, что $x_{k+1} \in U_{\mathbb{R}^n}(\widehat{x}, \delta_1) \cap M$.

Построим элемент x_{k+1}. Проверим сначала, что $y - G(x_k) \in V_2$. Действительно, $|y - G(x_k) - \Phi(\widehat{x})| \le |y - \Phi(\widehat{x}) - G(\widehat{x})| + |G(x_k) - G(\widehat{x})| = |y - F(\widehat{x})| + |G(x_k) - G(\widehat{x})| < \delta_2 + l|x_k - \widehat{x}| \le \delta_2 + l\delta_1$, т. е. $y - G(x_k) \in V_2$ в силу выбора δ_1 и δ_2.

Пусть элемент x_{k+1} такой, что

$$x_{k+1} \in \Phi^{-1}(y - G(x_k)) \tag{7}$$

и

$$|x_{k+1} - x_k| \le (1+\varepsilon)\mathrm{dist}\,(x_k, \Phi^{-1}(y - G(x_k)). \tag{8}$$

Теперь используя (8), метрическую регулярность Φ, равенство $\Phi(x_k) = y - G(x_{k-1})$, равносильное (7), липшивость G, а затем итерируя процедуру, будем иметь

$$|x_{k+1} - x_k| \le (1+\varepsilon)\mathrm{dist}\,(x_k, \Phi^{-1}(y - G(x_k)) \le (1+\varepsilon)c|y - G(x_k) \\ - \Phi(x_k)| = (1+\varepsilon)c|G(x_k) - G(x_{k-1})| \le (1+\varepsilon)cl|x_k - x_{k-1}| \\ = \alpha(\varepsilon)|x_k - x_{k-1}| \le \ldots \le \alpha^k(\varepsilon)|x_1 - x|. \quad (9)$$

Далее, используя неравенство треугольника, (9), формулу для суммы геометрической прогрессии, снова метрическую регулярность, выбор x и y, получим

$$|x_{k+1} - \widehat{x}| \le |x_{k+1} - x_k| + \ldots + |x_1 - \widehat{x}| \le (\alpha^k(\varepsilon) + \alpha^{k-1}(\varepsilon) \\ + \ldots + 1)|x_1 - x| + |x - \widehat{x}| < \frac{1+\varepsilon}{1-\alpha(\varepsilon)}\mathrm{dist}\,(x, \Phi^{-1}(y - G(x)) \\ + |x - \widehat{x}| \le \frac{(1+\varepsilon)c}{1-\alpha(\varepsilon)}|y - G(x) - \Phi(x)| + |x - \widehat{x}| \\ \le \frac{(1+\varepsilon)c}{1-\alpha(\varepsilon)}|y - F(\widehat{x})| + \frac{(1+\varepsilon)c}{1-\alpha(\varepsilon)}|F(x) - F(\widehat{x})| + |x - \widehat{x}| \\ < \frac{\delta_1}{3} + \frac{\delta_1}{3} + \frac{\delta_1}{3} = \delta_1, \quad (10)$$

т. е. $x_{k+1} \in U_{\mathbb{R}^n}(\widehat{x}, \delta_1)$ и значит, вся последовательность $\{x_k\}$ принадлежит $U_{\mathbb{R}^n}(\widehat{x}, \delta_1)$.

Последовательность $\{x_k\}$ фундаментальна. Действительно, для любых $k, l \in \mathbb{N}$, рассуждая точно также как и при доказательстве предыдущего неравенства, будем иметь

$$|x_{k+l} - x_k| \le |x_{k+l} - x_{k+l-1}| + \ldots + |x_{k+1} - x_k| \\ \le (\alpha^{k+l-1}(\varepsilon) + \ldots + \alpha^k(\varepsilon))|x_1 - x| < \frac{\alpha^k(\varepsilon)(1+\varepsilon)c}{1-\alpha(\varepsilon)}|y - F(x)|.$$

Положим $\varphi(x, y) = \lim_{k\to\infty} x_k$. Из (10) и замкнутости M следует, что $\varphi(x, y) \in U_{\mathbb{R}^n}(\widehat{x}, \delta)$. Включение (7) равносильно равенству $\Phi(x_k) = y - G(x_{k-1})$. Переходя здесь к пределу при $k \to \infty$, получаем, что $F(\varphi(x, y)) = y$.

Полагая в последнем неравенстве $k = 0$ и переходя к пределу при $l \to \infty$, приходим к неравенству $|\varphi(x, y) - x| \le ((1+\varepsilon)c/(1 - \alpha(\varepsilon)))|y - F(x)|$. Переходя здесь к пределу при $\varepsilon \to 0$, получаем, что $|\varphi(x, y) - x| \le (c/(1 - lc))|y - F(x)|$. Отсюда сразу следует, что $\mathrm{dist}(x, F^{-1}(y)) \le c'|y - F(x)|$ для всех $(x, y) \in (V_1' \cap M) \times V_2'$, где $c' = c/(1 - lc)$. $\square$

Следующее утверждение приводим без доказательства.

Лемма (о метрической регулярности линейного отображения). *Пусть* $\Lambda\colon \mathbb{R}^n \to \mathbb{R}^m$ *— линейный оператор,* K *— выпуклое множество в* $\mathbb{R}^n$, $\widehat{x} \in K$ *и* $0 \in \operatorname{int}\Lambda(K - \widehat{x})$. *Тогда отображение* $\Lambda|_K\colon K \to \mathbb{R}^m$, *где* $\Lambda|_K$ *— сужение* Λ *на* K, *метрически регулярно в окрестности точки* $\widehat{x}$.

Глава 2. Экстремальные и внутренние точки отображений

В этой главе мы сформулируем задачу о нахождении экстремальных точек отображения, которая является обобщением стандартной задачи о нахождении экстремальных точек функционала. Это позволит, с одной стороны, взглянуть на последнюю с более ясных геометрических позиций и получить для нее простые, но достаточно общего вида необходимые условия экстремума. С другой стороны, как будет показано в последующих двух главах, все рассматриваемые классы экстремальных задач можно весьма просто редуцировать к такой геометрической постановке, а получение необходимых условий экстремума для каждого класса будет сведено к расшифровке необходимых условий существование экстремальных точек соответствующего отображения.

§ 2.1. Основные определения. Связь со стандартной постановкой экстремальной задачи и метрической регулярностью.

Пусть M — непустое подмножество $\mathbb{R}^n$ и $F\colon M \to \mathbb{R}^m$.

Определение 2. *Назовем точку* $\widehat{x} \in M$ *экстремальной точкой отображения* F, *если найдется такая ее окрестность* V, *что* $F(\widehat{x}) \in \partial F(M \cap V)$, *и внутренней точкой отображения* F, *если* $F(\widehat{x}) \in \operatorname{int} F(M \cap V)$ *для любой окрестности* V *этой точки.*

Если $F(\widehat{x}) \in \partial F(M)$, *то будем говорить, что* $\widehat{x}$ *— глобально экстремальная точка отображения* F.

Заметим, что точка $\widehat{x}$ может быть либо внутренней, либо экстремальной точкой отображения F.

В последующих главах будет показано, что большинство классических задач может быть редуцировано к следующей постановке, где ограничения типа равенств выведены из общего множества ограничений:

$$f(x) \to \operatorname{extr}, \quad \mathcal{F}(x) = 0, \quad x \in M, \tag{11}$$

где $M \subset \mathbb{R}^n$, $f\colon M \to \mathbb{R}$ и $\mathcal{F}\colon M \to \mathbb{R}^m$.

Предложение 5. *Если* $\widehat{x}$ — *локальный* (*глобальный*) *экстремум в задаче* (11), *то* $\widehat{x}$ — *экстремальная* (*глобально экстремальная*) *точка отображения* $F = (f, \mathcal{F})\colon M \to \mathbb{R}^m$, *действующего по правилу* $F(x) = (f(x), \mathcal{F}(x))$, $x \in M$.

Доказательство. Пусть, для определенности, $\widehat{x}$ — локальный минимум. Тогда, по определению, найдется такая окрестность V точки $\widehat{x}$, что $f(x) \geq f(\widehat{x})$ для всех $x \in M \cap V$, для которых $\mathcal{F}(x) = 0$. Следовательно, для достаточно малых $\varepsilon > 0$ точка $(f(\widehat{x}) - \varepsilon, 0)$ не может принадлежать множеству $F(M \cap V)$ и поэтому $\widehat{x}$ — экстремальная точка отображения F.

В случае глобального экстремума доказательство аналогично. $\square$

В следующем параграфе, в достаточно общей ситуации, мы приведем необходимое условие (первого порядка) существования экстремальной точки отображения. Из этого условия единообразным способом будут получены необходимые условия экстремума в классических задачах математического и выпуклого программирования, вариационного исчисления и оптимального управления.

Двойственным к понятию экстремальной точки отображения является, в определенном смысле, понятие внутренней точки этого отображения. В следующем предложении приведены достаточные условия существования внутренней точки отображения.

Напомним, что понятие метрической регулярности отображения в окрестности данной точки было приведено в §1.5 (см. определение 1).

Предложение 6. *Если отображение* F *метрически регулярно в окрестности точки* $\widehat{x} \in M$, *то* $\widehat{x}$ — *внутренняя точка этого отображения.*

Доказательство. Пусть окрестности V_1, V_2 и константа c из неравенства (2), V — окрестность точки $\widehat{x}$, $r_1 > 0$ такое, что $U_{\mathbb{R}^n}(\widehat{x}, r_1) \subset V$, а $0 < r_2 \leq r_1/2c$ таково, что $U_{\mathbb{R}^m}(F(\widehat{x}), r_2) \subset V_2$. Покажем, что $U_{\mathbb{R}^m}(F(\widehat{x}), r_2) \subset F(M \cap V)$ и тем самым докажем предложение. Пусть $y \in U_{\mathbb{R}^m}(F(\widehat{x}), r_2)$. Тогда из (2), при $x = \widehat{x}$, следует, что существует $x(y) \in F^{-1}(y)$, для которого $|\widehat{x} - x(y)| \leq 2\,\mathrm{dist}\,(\widehat{x}, F^{-1}(y)) \leq 2c|y - F(\widehat{x})| < 2cr_2 \leq r_1$, т. е. $x(y) \in F^{-1}(y) \cap U_{\mathbb{R}^n}(\widehat{x}, r_1)$. Таким образом, $y \in F(M \cap U_{\mathbb{R}^n}(\widehat{x}, r_1)) \subset F(M \cap V)$. $\square$

§ 2.2. Основная теорема и ее следствия.

Необходимые и достаточные условия глобально экстремальной точки выпуклого отображения. Основная теорема. Теорема Люстерника.

Пусть по прежнему M — непустое подмножество $\mathbb{R}^n$ и $F\colon M \to \mathbb{R}^m$. Мы начинаем с наиболее простой ситуации, когда F — выпуклое отображение.

Скажем, что отображение F *выпукло*, если множество $F(M)$ выпукло.

Напомним, что понятие нормального конуса ко множеству в данной точке было определено в §1.3.

Предложение 7. *Пусть F — выпуклое отображение. Для того, что бы точка $\widehat{x} \in M$ была глобально экстремальной точкой для F необходимо и достаточно, чтобы существовал такой ненулевой вектор $\lambda \in (\mathbb{R}^m)^*$, что*

$$\langle \lambda, F(x) - F(\widehat{x}) \rangle \geq 0, \qquad \forall\, x \in M. \tag{12}$$

Доказательство. В силу определения нормального конуса выполнение соотношения (12) с ненулевым вектором λ означает, что $N_{F(M)}(F(\widehat{x})) \neq \{0\}$. По определению $F(\widehat{x})$ — граничная точка выпуклого множества $F(M)$, поэтому утверждение данного предложения сразу следует из предложения 2 §1.3. с заменой A на $F(M)$ и $\widehat{x}$ на $F(\widehat{x})$. $\square$

Далее будем предполагать, что M — выпуклое замкнутое множество в $\mathbb{R}^n$ и писать K вместо M, так что будем иметь дело с отображением $F\colon K \to \mathbb{R}^m$.

Пусть $\widehat{x} \in K$ и отображение F дифференцируемо в $\widehat{x}$ относительно K. Для $\lambda \in (\mathbb{R}^m)^*$ рассмотрим включение

$$-(F'(\widehat{x}))^*\lambda \in N_K(\widehat{x}), \tag{13}$$

которое, очевидно, равносильно неравенству

$$\langle \lambda, F'(\widehat{x})(x - \widehat{x} \rangle \geq 0, \qquad \forall\, x \in K. \tag{14}$$

Ясно, что если λ удовлетворяет (13) ((14)), то вектор $\alpha\lambda$, где $\alpha \geq 0$, также удовлетворяет этим соотношениям, поэтому, если $\lambda \neq 0$, то будем считать, что $|\lambda| = 1$.

Обозначим через $\Lambda(\widehat{x})$ совокупность тех $\lambda \in (\mathbb{R}^m)^*$, $|\lambda| = 1$, для которых справедливо включение (13) (или неравенство (14)).

Теорема (Основная теорема)**.** *Пусть K — выпуклое замкнутое подмножество $\mathbb{R}^n$ и отображение $F\colon K \to \mathbb{R}^m$ строго дифференцируемо в точке $\widehat{x} \in K$ относительно K. Тогда, если $\widehat{x}$ — экстремальная точка отображения F, то $\Lambda(\widehat{x}) \neq \emptyset$. С другой стороны,* *если* $\Lambda(\widehat{x}) = \emptyset$*, то* $\widehat{x}$ *— внутренняя точка отображения* F*, и более того,* F *метрически регулярно в окрестности* $\widehat{x}$*.*

Доказательство. Для доказательства теоремы достаточно показать, что если $\Lambda(\widehat{x}) = \emptyset$, то отображение F метрически регулярно в окрестности точки $\widehat{x}$. Действительно, если это так, то из предложения 6 следует, что $\widehat{x}$ — внутренняя этого отображения. Следовательно, если $\widehat{x}$ — не внутренняя точка F, а значит, экстремальная точка F, то $\Lambda(\widehat{x}) \neq \emptyset$.

Итак, пусть $\Lambda(\widehat{x}) = \emptyset$. Тогда $0 \in \operatorname{int} F'(\widehat{x})(K - \widehat{x})$. Действительно, если это включение не выполняется, или равносильно, $F'(\widehat{x})\widehat{x} \notin \operatorname{int} F'(\widehat{x})(K)$, то $\widehat{x}$ — глобально экстремальная точка для выпуклого отображения $x \mapsto F'(\widehat{x})x$ на K (образ выпуклого множества при линейном отображении — выпуклое множество). Тогда в силу предложения 7 найдется ненулевой вектор $\lambda \in (\mathbb{R}^m)^*$ (можно считать $|\lambda| = 1$), для которого справедливо соотношение (12) и которое в данном случае есть неравенство (14), а значит, $\Lambda(\widehat{x}) \neq \emptyset$.

Положим $\Phi = F'(\widehat{x})|_K$ и $G = F - \Phi$. Поскольку $0 \in \operatorname{int} F'(\widehat{x})(K - \widehat{x})$, то по лемме о метрической регулярности линейного отображения Φ метрически регулярно в окрестности точки $\widehat{x}$ с некоторой константой $c > 0$. Пусть $0 < l < c$. Так как F строго дифференцируемо в точке $\widehat{x}$ относительно K, то найдется такая окрестность V_0 точки $\widehat{x}$, что отображение G будет липшицево на $V_0 \cap K$ с константой Липшица l. Но тогда по теореме о возмущении отображение $F = \Phi + G$ будет метрически регулярно в окрестности точки $\widehat{x}$. $\square$

Пусть M — непустое подмножество $\mathbb{R}^n$. Элемент $h \in \mathbb{R}^n$ называется *касательным вектором к* M *в точке* $\widehat{x} \in M$, если существуют $\varepsilon > 0$ и отображение $r\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ такие, что $\widehat{x} + th + r(t) \in M$ для всех $t \in (-\varepsilon, \varepsilon)$, где $r(t) = o(t)$ при $t \to 0$. Множество всех касательных векторов к M в точке $\widehat{x} \in M$ обозначим $T_{\widehat{x}}M$.

Если U — окрестность точки $\widehat{x} \in \mathbb{R}^n$ и $F\colon U \to \mathbb{R}^m$, то положим $M(\widehat{x}) = \{\, x \in U : F(x) = F(\widehat{x}) \,\}$, т. е. $M(\widehat{x})$ — "поверхность уровня" отображения F.

Следствие 1 (Теорема Люстерника)**.** *Пусть* U *— окрестность точки* $\widehat{x} \in \mathbb{R}^n$*, отображение* $F\colon U \to \mathbb{R}^m$ *строго дифференцируемо в* $\widehat{x}$ *и* $\operatorname{Im} F'(\widehat{x}) = \mathbb{R}^m$*. Тогда*

$$T_{\widehat{x}}M(\widehat{x}) = \operatorname{Ker} F'(\widehat{x}).$$

Доказательство. Докажем, что $T_{\widehat{x}}M(\widehat{x}) \subset \operatorname{Ker} F'(\widehat{x})$. Пусть $h \in T_{\widehat{x}}M$ и отображение 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 $t \to 0$, получаем, что $h \in \operatorname{Ker} F'(\widehat{x})$.

Докажем теперь, что $\operatorname{Ker} F'(\widehat{x}) \subset T_{\widehat{x}} M(\widehat{x})$. Сначала заметим, что так как $\operatorname{Im} F'(\widehat{x}) = \mathbb{R}^m$, то неравенство $\langle \lambda, F'(\widehat{x})(x - \widehat{x}) \rangle \geq 0$ для всех $x \in \mathbb{R}^n$ возможно только при $\lambda = 0$, т. е. $\Lambda(\widehat{x}) = \emptyset$ и значит, отображение F метрически регулярно в окрестности точки $\widehat{x}$. Следовательно, найдется окрестность $V \subset U$ точки $\widehat{x}$ и константа $c > 0$ такие, что для любого $x \in V$ (учитывая, что $F^{-1}(F(\widehat{x})) = M(\widehat{x})$) справедливо оценка $\operatorname{dist}(x, M(\widehat{x})) \leq c|F(x) - F(\widehat{x})|$.

Пусть $h \in \operatorname{Ker} F'(\widehat{x})$. Тогда существует $\varepsilon > 0$ такое, что $\widehat{x} + th \in V$ для всех $t \in (-\varepsilon, \varepsilon)$. Из оценки выше следует, что для каждого такого t найдется элемент $x(t) \in M(\widehat{x})$, для которого $|\widehat{x} + th - x(t)| \leq 2\operatorname{dist}(\widehat{x} + th, M(\widehat{x})) \leq 2c|F(\widehat{x} + th) - F(\widehat{x})| = 2c|F(\widehat{x}) + tF'(\widehat{x})h + o(t) - F(\widehat{x})| = |o(t)|$. Если обозначить $r(t) = x(t) - \widehat{x} - th$, то получим, что $\widehat{x} + th + r(t) \in M(\widehat{x})$ для всех $t \in (-\varepsilon, \varepsilon)$ и $r(t) = o(t)$ при $t \to 0$, т. е. $h \in T_{\widehat{x}} M(\widehat{x})$. $\square$

Глава 3. Условия экстремума в задачах математического и выпуклого программирования

§ 3.1. Гладкие задачи без ограничений.

Теорема Ферма. Необходимые и достаточные условия экстремума второго порядка.

Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$ и $f\colon U \to \mathbb{R}$. Рассмотрим задачу

$$f(x) \to \operatorname{extr}, \quad x \in U. \tag{P_1}$$

Теорема (Ферма). *Пусть $\widehat{x}$ — локальный экстремум в задаче (P_1). Если функция f строго дифференцируема в точке $\widehat{x}$, то $f'(\widehat{x}) = 0$.*

Доказательство. Пусть $r > 0$ такое, что $B_{\mathbb{R}^n}(\widehat{x}, r) \subset U$. Очевидно, что $\widehat{x}$ будет локальным экстремумом и в задаче

$$f(x) \to \operatorname{extr}, \quad x \in B_{\mathbb{R}^n}(\widehat{x}, r),$$

которая имеет вид задачи (11), где $K = B_{\mathbb{R}^n}(\widehat{x}, r)$, а ограничение $\mathcal{F}(x) = 0$ отсутствует. Таким образом $F = f$ и согласно предложению 5 §2.1 точка $\widehat{x}$ будет экстремальной для функционала f. В силу Основной теоремы найдется число λ такое, что $|\lambda| = 1$ и $-(f'(\widehat{x}))^* \lambda = -\lambda f'(\widehat{x}) \in N_K(\widehat{x})$. Но $\widehat{x}$ — внутренняя точка K и поэтому $N_K(\widehat{x}) = \{0\}$ в силу предложения 7 §2.2 и значит, $f'(\widehat{x}) = 0$. $\square$

Как уже говорилось, вывод необходимых условий экстремума (первого порядка) во всех задачах будет проводиться по единой схеме, опираясь на Основную теорему. В рассмотренной простейшей ситуации можно, конечно, привести и непосредственное доказательство, предполагающее лишь дифференцируемость функции f в точке $\widehat{x}$ (проделайте это). Но в последующих примерах будет видно, что предлагаемый подход оказывается весьма эффективен.

Теорема (Необходимые и достаточные условия экстремума второго порядка для задачи (P_1))**.** *Пусть в задаче (P_1) функция f дважды дифференцируема в точке $\widehat{x}$. Тогда*

1) *если $\widehat{x}$ — локальный минимум (максимум) функции f, то для любого $h \in \mathbb{R}^n$ выполняется неравенство*

$$f''(\widehat{x})[h,h] \geq 0 \ \ (\leq 0);$$

2) *если $f'(\widehat{x}) = 0$ и для любого $h \in \mathbb{R}^n$, $h \neq 0$, выполняется неравенство*

$$f''(\widehat{x})[h,h] > 0 \ \ (< 0),$$

то $\widehat{x}$ — локальный минимум (максимум) функции f.

Доказательство. 1) Пусть, для определенности, $\widehat{x}$ — локальный минимум функции f и $h \in \mathbb{R}^n$. По теорема Ферма $f'(\widehat{x}) = 0$ и тогда по формуле Тейлора для достаточно малых $t \in \mathbb{R}$ имеем $0 \leq f(\widehat{x}+th) - f(\widehat{x}) = (t^2/2)f''(\widehat{x})[h,h] + o(t^2)$. Отсюда, деля обе части неравенства на t^2 и устремляя t к нулю, получаем требуемое. Для локального максимума рассуждения аналогичны.

2) Пусть $f''(\widehat{x})[h,h] > 0$ для любого ненулевого $h \in \mathbb{R}^n$. Функция $h \to f''(\widehat{x})[h,h]$ непрерывна на $\mathbb{R}^n$. Обозначим через m ее минимальное значение на единичной сфере $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$. Ясно, что $m > 0$. Снова по формуле Тейлора имеем для любого $h \neq 0$ такого, что $\widehat{x} + h \in U$ (учитывая, что $p = |h|^{-1}h \in S^{n-1}$)

$$f(\widehat{x}+h) = f(\widehat{x}) + \frac{1}{2}f''(\widehat{x})[h,h] + o(|h|^2) = f(\widehat{x})$$
$$+ \frac{|h|^2}{2}\left(f''(\widehat{x})[p,p] + \frac{o(|h|^2)}{|h|^2}\right) \geq f(\widehat{x}) + \frac{|h|^2}{2}\left(m + \frac{o(|h|^2)}{|h|^2}\right).$$

Выражение в скобках в правой части неравенства положительно для достаточно малых h. Следовательно, $f(\widehat{x}+h) > f(\widehat{x})$ для таких h и значит, $\widehat{x}$ — локальный минимум функции f. Для локально максимума рассуждения аналогичны. $\square$

§ 3.2. Гладкие задачи с ограничениями типа равенств.

3.2.1. Правило множителей Лагранжа — необходимое условие экстремума первого порядка.

Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$ и $f_i \colon U \to \mathbb{R}$, $i = 0, 1, \ldots, m$. Задача

$$f_0(x) \to \mathrm{extr}, \quad f_i(x) = 0, \quad i = 1, \ldots, m \qquad (P_2)$$

называется *задачей с ограничениями, задаваемые равенствами* или *задачей с ограничениями типа равенств.*

Свяжем с задачей (P_2) следующую функцию Лагранжа: $\mathcal{L} \colon U \times (\mathbb{R}^{m+1})^*$, $\mathcal{L}(x, \lambda) = \sum_{i=0}^m \lambda_i f_i(x)$, где $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ — набор множителей Лагранжа.

Теорема (Правило множителей Лагранжа — необходимое условие экстремума первого порядка). *Пусть $\widehat{x}$ — локальный экстремум в задаче (P_2). Если функции f_i, $i = 0, 1, \ldots, m$, строго дифференцируемы в точке $\widehat{x}$, то найдется такой ненулевой набор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, что*

$$\mathcal{L}_x(\widehat{x}, \lambda) = 0 \;\Leftrightarrow\; \sum_{i=0}^m \lambda_i f_i'(\widehat{x}) = 0. \qquad (15)$$

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

Доказательство. Пусть $r > 0$ такое, что $B_{\mathbb{R}^n}(\widehat{x}, r) \subset U$. Согласно предложению 5 §2.1 точка $\widehat{x}$ будет экстремальной для отображения $F \colon K \to \mathbb{R}^{m+1}$, где $F(x) = (f_0(x), f_1(x), \ldots, f_m(x))^T$, а $K = B_{\mathbb{R}^n}(\widehat{x}, r)$. В силу Основной теоремы найдется вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$, $|\lambda| = 1$, такой, что $(F'(\widehat{x}))^* \lambda \in N_K(\widehat{x})$. Тогда, используя формулу для сопряженного оператора из §1.1, формулу для производной отображения F, задаваемого набором функций и то, что $\widehat{x}$ — внутренняя точка K, будем иметь

$$(F'(\widehat{x}))^* \lambda = \lambda F'(\widehat{x}) = \sum_{i=0}^m \lambda_i f_i'(\widehat{x}) \in N_K(\widehat{x}) = \{0\}.$$

Последнее утверждение теоремы непосредственно следует из (15) и определения линейной независимости векторов. □

3.2.2. Условия экстремума второго порядка.

Необходимые условия второго порядка. Достаточные условия второго порядка.

Теорема (Необходимые условия экстремума второго порядка для задачи (P_2))**.** *Пусть* $\widehat{x}$ *— локальный минимум (максимум) в задаче* (P_2)*. Тогда, если функции* f_i*,* $i = 0, 1, \ldots, m$*, дважды дифференцируемы в точке* $\widehat{x}$ *и векторы* $f_1'(\widehat{x}), \ldots, f_m'(\widehat{x})$ *линейно независимы, то существует такой набор множителей Лагранжа* $\lambda = (1, \lambda_1, \ldots, \lambda_m)$*, что*

$$\mathcal{L}_x(\widehat{x}, \lambda) = 0 \iff f_0'(\widehat{x}) + \sum_{i=1}^{m} \lambda_i f_i'(\widehat{x}) = 0$$

и для любого $h \in \mathbb{R}^n$*, для которого* $\langle f_i'(\widehat{x}), h \rangle = 0$*,* $i = 1, \ldots, m$*, справедливы соотношения*

$$\mathcal{L}_{xx}(\widehat{x}, \lambda)[h, h] \geq 0 \ (\leq 0) \iff f_0''(\widehat{x})[h, h] + \sum_{i=1}^{m} \lambda_i f_i''(\widehat{x})[h, h] \geq 0 \ (\leq 0).$$

Доказательство. Воспользуемся теоремой Люстерника. Для этого определим отображение $F \colon U \to \mathbb{R}^m$ по правилу $F(x) = (\, f_1(x), \ldots, f_m(x)\,)^T$. Так как функции f_i, $i = 1, \ldots, m$, дважды дифференцируемы в $\widehat{x}$, то их производные непрерывны в этой точке и значит, строго дифференцируемы в $\widehat{x}$, а тогда отображение F строго дифференцируемо в $\widehat{x}$. Далее, $F'(\widehat{x})(\mathbb{R}^n) = \mathbb{R}^m$, поскольку векторы $f_1'(\widehat{x}), \ldots, f_m'(\widehat{x})$ линейно независимы. Условие $\langle f_i'(\widehat{x}), h \rangle = 0$, $i = 1, \ldots, m$, равносильно тому, что $h \in \operatorname{Ker} F'(\widehat{x})$. Следовательно, по теореме Люстерника, h — касательный вектор ко множеству $\{\, x \in \mathbb{R}^n : F(x) = F(\widehat{x}) = 0 \,\}$, т. е. существуют $\varepsilon > 0$ и отображение $r \colon (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ такие, что $F(\widehat{x} + th + r(t)) = 0$ (или $f_i(\widehat{x} + th + r(t)) = 0$, $i = 1, \ldots, m$) для $t \in (-\varepsilon, \varepsilon)$, где $r(t) = o(t)$ при $t \to 0$.

Таким образом, для малых t точки $\widehat{x} + th + r(t)$ допустимы в задаче (P_2). Так как $\widehat{x}$ — локальный экстремум, то согласно правилу множителей Лагранжа найдется ненулевой набор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ такой, что $\mathcal{L}_x(\widehat{x}, \lambda) = 0$, причем $\lambda_0 \neq 0$ и можно считать, что $\lambda_0 = 1$. Учитывая это, будем иметь для достаточно малых t по формуле Тейлора, считая для определенности, что $\widehat{x}$ — локальный минимум:

$$0 \leq f_0(\widehat{x} + th + r(t)) - f_0(\widehat{x}) = \mathcal{L}(\widehat{x} + th + r(t), \lambda) - \mathcal{L}(\widehat{x}, \lambda)$$
$$= \frac{1}{2}\,\mathcal{L}_{xx}(\widehat{x}, \lambda)[th + r(t), th + r(t)] + o(t^2) = \frac{t^2}{2}\mathcal{L}_{xx}(\widehat{x}, \lambda)[h, h] + o(t^2),$$

откуда, деля на t^2 и переходя к пределу при $t \to \infty$, получаем нужное утверждение.

Для случая локального максимума рассуждения аналогичны. $\square$

Теорема (Достаточные условия экстремума второго порядка для задачи (P_2))**.** *Пусть в задаче (P_2) функции f_i, $i = 0, 1, \dots, m$, дважды дифференцируемы в допустимой точке $\widehat{x}$. Тогда, если существует такой набор множителей Лагранжа $\lambda = (1, \lambda_1, \dots, \lambda_m)$, что*

$$\mathcal{L}_x(\widehat{x}, \lambda) = 0$$

и для любого ненулевого $h \in \mathbb{R}^n$, для которого $\langle f_i'(\widehat{x}), h\rangle = 0$, $i = 1, \dots, m$, справедливы соотношения

$$\mathcal{L}_{xx}(\widehat{x}, \lambda)[h, h] > 0 \ \ (< 0),$$

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

Доказательство. Доказываем от противного. Пусть $\widehat{x}$ не является локальным экстремумом в задаче (P_2). Для определенности считаем, что $\widehat{x}$ — не локальный минимум. Покажем, что в этом случае найдется такое ненулевое $\overline{h} \in \mathbb{R}^n$, что $\langle f_i'(\widehat{x}), \overline{h}\rangle = 0$, $i = 1, \dots, m$, и для любого набора $\lambda = (1, \lambda_1, \dots, \lambda_m)$, для которого $\mathcal{L}_x(\widehat{x}, \lambda) = 0$, выполняется неравенство

$$\mathcal{L}_{xx}(\widehat{x}, \lambda)[\overline{h}, \overline{h}] \le 0. \tag{16}$$

Действительно, так как $\widehat{x}$ — не локальный минимум, то существует последовательность $\{x_k\}_{k\in\mathbb{N}}$ допустимых элементов в (P_2) такая, что $x_k \neq \widehat{x}$, $x_k \to \widehat{x}$ при $k \to \infty$ и $f_0(x_k) < f_0(\widehat{x})$. Обозначим $h_k = x_k - \widehat{x}$ и $p_k = |h_k|^{-1}|h_k|$.

Пусть $\lambda = (1, \lambda_1, \dots, \lambda_m)$ и $\mathcal{L}_x(\widehat{x}, \lambda) = 0$. Учитывая, что $f_i(x_k) = f_i(\widehat{x}){=}0$, $i = 1, \dots, m$, будем иметь по формуле Тейлора

$$\begin{aligned} f_0(x_k) - f_0(\widehat{x}) = \mathcal{L}(x_k, \lambda) - \mathcal{L}(\widehat{x}, \lambda) = \langle \mathcal{L}_x(\widehat{x}, \lambda), h_k\rangle \\ + \frac{|h_k|^2}{2}\mathcal{L}_{xx}(\widehat{x}, \lambda)[p_k, p_k] + o(|h_k|^2). \end{aligned} \tag{17}$$

Выражение слева отрицательно, а первое слагаемое справа равно нулю по условию. Деля (17) на $|h_k|^2/2$, приходим к неравенству

$$\mathcal{L}_{xx}(\widehat{x}, \lambda)[p_k, p_k] + \frac{o(|h_k|^2)}{|h_k|^2} < 0. \tag{18}$$

Последовательность $\{p_k\}$ принадлежит единичной сфере в $\mathbb{R}^n$ и поэтому из нее можно выделить подпоследовательность, сходящуюся к некоторому $\overline{h} \in \mathbb{R}^n$, $|\overline{h}| = 1$. Будем считать, что сама последовательность сходится к $\overline{h}$. Переходя к пределу при $k \to \infty$ в (18), получаем неравенство (16). Осталось проверить, что $\langle f_i'(\widehat{x}), \overline{h}\rangle = 0$, $1 \le i \le m$. Для всех $k \in \mathbb{N}$ и $1 \le i \le m$ имеем

$$f_i(x_k) = f_i(\widehat{x}) + \langle f_i'(\widehat{x}), h_k\rangle + o(h_k) \;\Leftrightarrow\; \langle f_i'(\widehat{x}), h_k\rangle + o(h_k) = 0.$$

Деля каждое равенство на $|h_k|$ и переходя к пределу, получаем, что $\langle f_i'(\widehat{x}), \overline{h}\rangle = 0$, $i = 1, \ldots, m$. □

§ 3.3. Гладкие задачи с ограничениями вида $G(x) \in Q$.

3.3.1. Правило множителей Лагранжа — необходимое условие экстремума первого порядка.

Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$, $f_0\colon U \to \mathbb{R}$, $G\colon U \to \mathbb{R}^m$ и Q — непустое подмножество $\mathbb{R}^m$. Рассмотрим экстремальную задачу

$$f_0(x) \to \min, \quad G(x) \in Q. \tag{P_3}$$

Сопоставим этой задаче следующую функцию Лагранжа

$$\mathcal{L}(x, \lambda) = \lambda_0 f_0(x) + \langle \overline{\lambda}, G(x)\rangle,$$

где $\lambda = (\lambda_0, \overline{\lambda}) \in \mathbb{R} \times (\mathbb{R}^m)^*$.

Теорема (Правило множителей Лагранжа — необходимые условия минимума первого порядка в задаче (P_3))**.** *Пусть $\widehat{x}$ — локальный минимум в задаче (P_3). Если функция f_0 и отображение G строго дифференцируемы в точке $\widehat{x}$, а множество Q выпукло и замкнуто, то найдется такой ненулевой набор множителей Лагранжа $\lambda = (\lambda_0, \overline{\lambda}) \in \mathbb{R} \times (\mathbb{R}^m)^*$, что*

$$\mathcal{L}_x(\widehat{x}, \lambda) = 0 \;\Leftrightarrow\; \lambda_0 f_0'(\widehat{x}) + (G'(\widehat{x}))^* \overline{\lambda} = 0 \tag{19}$$

и

$$\lambda_0 \ge 0, \quad \overline{\lambda} \in N_Q(G(\widehat{x})). \tag{20}$$

Доказательство. Сведем задачу (P_3) к задаче на исследование экстремальной точки отображения, определенного на выпуклом множестве. Пусть $r > 0$ такое, что $B_{\mathbb{R}^n}(\widehat{x}, r) \subset U$. Введем новую переменную $z = (x, u_0, u)^T \in B_{\mathbb{R}^n}(\widehat{x}, r) \times \mathbb{R}_+ \times Q = K$ и положим $f(z) = f_0(x) + u_0$ и $\mathcal{F}(z) = G(x) - u$.

Рассмотрим задачу

$$f(z) \to \min, \quad \mathcal{F}(z) = 0, \quad z \in K.$$

Легко видеть, что если $\widehat{x}$ является локальным минимумом в задаче (P_3), то точка $\widehat{z} = (\widehat{x}, 0, G(\widehat{x}))^T$ — локальный минимум в данной задаче. Согласно предложению 5 §2.1 точка $\widehat{x}$ будет экстремальной для отображения $F = (f, \mathcal{F})\colon K \to \mathbb{R}^{m+1}$. В силу Основной теоремы найдется вектор $\lambda \in (\mathbb{R}^{m+1})^*$, $|\lambda| = 1$, такой, что $\langle \lambda, F'(\widehat{z})(z-\widehat{z})\rangle \geq 0$ для всех $z \in K$. Далее, используя последовательно предложения 3, 1 и 4 из гл. 1, будем иметь (учитывая, что $f_{u_0}(\widehat{z}) = 1$, $\mathcal{F}_x(\widehat{z}) = G'(\widehat{x})$ и $\mathcal{F}_u(\widehat{z}) = -\mathrm{Id}$, где Id — тождественный оператор)

$$\begin{aligned} 0 \leq \langle \lambda, F'(\widehat{z})(z-\widehat{z})\rangle = \langle \lambda, (f'(\widehat{z}), \mathcal{F}'(\widehat{z}))(z-\widehat{z})\rangle = \lambda_0 \langle f'(\widehat{z}), z-\widehat{z}\rangle \\ + \langle \overline{\lambda}, \mathcal{F}'(\widehat{z})(z-\widehat{z})\rangle = \lambda_0(\langle f_0'(\widehat{x}), x-\widehat{x}\rangle + u_0) + \langle \overline{\lambda}, G'(\widehat{x})(x-\widehat{x}) \\ -(u-G(\widehat{x})\rangle = \langle \lambda_0 f_0'(\widehat{x}) + (G'(\widehat{x}))^*\overline{\lambda}, x-\widehat{x}\rangle + \lambda_0 u_0 \\ -\langle \overline{\lambda}, u-G(\widehat{x})\rangle \quad (21) \end{aligned}$$

для всех $x \in B_{\mathbb{R}^n}(\widehat{x}, r)$, $u_0 \geq 0$ и $u \in Q$.

Полагая здесь $x = \widehat{x}$ и $u = G(\widehat{x})$, получаем, что $\lambda_0 \geq 0$.

Пусть в (21) $u_0 = 0$ и $u = G(\widehat{x})$, тогда $\langle \lambda_0 f_0'(\widehat{x}) + (G'(\widehat{x}))^*\overline{\lambda}, x-\widehat{x}\rangle \geq$ для всех $x \in B_{\mathbb{R}^n}(\widehat{x}, r)$. Поскольку $\widehat{x}$ — внутренняя точка шара, то $\lambda_0 f_0'(\widehat{x}) + (G'(\widehat{x}))^*\overline{\lambda} = 0$.

Наконец, полагая $x = \widehat{x}$ и $u_0 = 0$, получаем, что $\overline{\lambda} \in N_Q(G(\widehat{x}))$. $\square$

Упражнение. *Доказать, что если выполнено условие регулярности:* $0 \in \operatorname{int}(\operatorname{Im} G'(\widehat{x}) + G(\widehat{x}) - Q)$, *то* $\lambda_0 \neq 0$.

3.3.2. Следствие: гладкие задачи с ограничениями типа равенств и неравенств.

Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$ и $f_i \colon U \to \mathbb{R}$, $i = 0, 1, \ldots, m$. Задача

$$f_0(x) \to \min, \quad f_i(x) \leq 0, \quad 1 \leq i \leq m', \quad f_i(x) = 0, \\ m'+1 \leq i \leq m \quad (P_4)$$

называется *задачей с ограничениями, задаваемые равенствами и неравенствами* или *задачей с ограничениями типа равенств и неравенств*.

Свяжем с задачей (P_4) функцию Лагранжа $\mathcal{L} \colon U \times (\mathbb{R}^{m+1})^* \to \mathbb{R}$, $\mathcal{L}(x, \lambda) = \sum_{i=0}^m \lambda_i f_i(x)$, которая, как мы видим, имеет тот же вид, что и в задаче (P_2).

Теорема (Правило множителей Лагранжа — необходимые условия минимума первого порядка)**.** *Если точка* $\widehat{x} \in U$ *является локальным минимумом в задаче* (P_4) *и функции* f_i, $i = 0, 1, \ldots, m$, *строго дифференцируемы в* $\widehat{x}$, *то найдется такой ненулевой набор множителей Лагранжа* $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, *что*

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

(b) $\lambda_i \geq 0,\ i=0,1,\ldots,m'$;

(c) $\lambda_i f_i(\widehat{x})=0,\ i=1,\ldots,m'$.

Доказательство. Если определить отображение $G\colon U\to\mathbb{R}^m$ по формуле $G(x)=(f_1(x),\ldots,f_m(x))^T$ и положить $Q=\mathbb{R}_-^{m'}\times\{0\}$, где 0 — нулевой элемент в $\mathbb{R}^{m-m'}$, то задача (P_4) примет в точности вид задачи (P_3). Используя представление для сопряженного оператора из §1.1, условие (19), где $\overline{\lambda}=(\lambda_1,\ldots,\lambda_m)$, перепишется так $0=\lambda_0 f_0'(\widehat{x})+(G'(\widehat{x}))^*\overline{\lambda}=\lambda_0 f_0'(\widehat{x})+\overline{\lambda}G'(\widehat{x})=\lambda_0 f_0'(\widehat{x})+\sum_{i=1}^m \lambda_i f_i'(\widehat{x})=\sum_{i=0}^m \lambda_i f_i'(\widehat{x})$, т. е. соотношение (a) доказано и при этом, $\lambda_0\geq 0$.

Включение $\overline{\lambda}\in N_Q(G(\widehat{x}))$ в нашем случае означает, что

$$\langle\overline{\lambda},u-G(\widehat{x})\rangle=\sum_{i=1}^m \lambda_i(u_i-f_i(\widehat{x}))=\sum_{i=1}^{m'}\lambda_i(u_i-f_i(\widehat{x}))\leq 0,\qquad \forall\, u_i\leq 0,\ \ i=1,\ldots,m'. \tag{22}$$

Пусть $1\leq i_0\leq m'$. Подставляя в это неравенство $u_{i_0}=f_{i_0}(\widehat{x})-1$ и $u_i=f_i(\widehat{x})$, если $i\neq i_0$, получаем, что $\lambda_{i_0}\geq 0$. Этим доказано соотношение (b). Отсюда следует, что $\lambda_i f_i(\widehat{x})\leq 0$, $i=1,\ldots,m'$. Но подставляя в (22) $u_1=\ldots=u_{m'}=0$, получим, что $\sum_{i=1}^{m'}\lambda_i f_i(\widehat{x})\geq 0$. Следовательно, $\lambda_i f_i(\widehat{x})=0$, $i=1,\ldots,m'$, что и есть утверждение (c). □

Упражнение. *Доказать, что если векторы $f_i'(\widehat{x})$, $i=m'+1,\ldots,m$, линейно независимы и существует вектор $h\in\mathbb{R}^n$ такой, что $\langle f_i'(\widehat{x}),h\rangle<0$ для тех индексов $1\leq i\leq m'$, для которых $f_i(\widehat{x})=0$ и $\langle f_i'(\widehat{x}),h\rangle=0$ при $m'+1\leq i\leq m$, то $\lambda_0\neq 0$.*

§ 3.4. Выпуклые задачи.

Выпуклые функции. Неравенство Иенссена. Теорема Каруша-Куна-Таккера.

Пусть Q — выпуклое подмножество $\mathbb{R}^n$. Функция $f\colon Q\to\mathbb{R}$ называется *выпуклой*, если множество

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

которое называется *надграфиком* f, выпукло.

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

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

которое называется *неравенством Иенссена.*

Пусть $f_i\colon Q \to \mathbb{R}$, $i = 0, 1, \ldots, m$ — выпуклые функции. Задача

$$f_0(x) \to \min, \quad f_i(x) \le 0, \quad 1 \le i \le m, \quad x \in Q \qquad (P_5)$$

называется *выпуклой задачей* или *задачей выпуклого программирования.*

Заметим, что в этой задаче любой локальный минимум является и глобальным. Действительно, пусть $\widehat{x}$ — локальный минимум, т. е. существует такая окрестность U точки $\widehat{x}$, что $f_0(\widehat{x}) \le f_0(x)$ для всех допустимых $x \in U$. Пусть x — произвольная допустимая точка. Для достаточно малых $0 < \alpha \le 1$ точки $(1-\alpha)\widehat{x} + \alpha x = \widehat{x} + \alpha(x - \widehat{x})$ принадлежат U и допустимы, так как $(1-\alpha)\widehat{x} + \alpha x \in Q$, и по неравенству Иенссена $f_i(1-\alpha)\widehat{x} + \alpha x) \le (1-\alpha)f_i(\widehat{x}) + \alpha f_i(x) \le 0$, $i = 1, \ldots, m$. Снова по неравенству Иенссена, $f_0(\widehat{x}) \le f_0((1-\alpha)\widehat{x} + \alpha x) \le (1-\alpha)f_0(\widehat{x}) + \alpha f_0(x)$, т. е. $f_0(\widehat{x}) \le f_0(x)$.

С задачей (P_5) связжем функция Лагранжа $\mathcal{L}\colon X \times (\mathbb{R}^{m+1})^* \to \mathbb{R}$, $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^{m} \lambda_i f_i(x)$, которая имеет тот же вид, что и в задачах (P_2) и (P_4).

Теорема (Каруша–Куна–Таккера)**.** *Если $\widehat{x}$ — минимум в задаче (P_5), то найдется такой ненулевой набор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, что*

(*a*) $\min\limits_{x \in Q} \mathcal{L}(x, \lambda) = \mathcal{L}(\widehat{x}, \lambda)$;

(*b*) $\lambda_i \ge 0, \quad i = 0, 1, \ldots, m$;

(*c*) $\lambda_i f_i(\widehat{x}) = 0, \quad i = 1, \ldots, m.$

Если существует допустимая в (P_5) точка $\widehat{x}$ и набор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$, удовлетворяющие условиям (a), (b) и (c) и при этом $\lambda_0 > 0$, то $\widehat{x}$ — решение задачи (P_5).

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

Доказательство. Введем новую переменную $z = (x, u)^T \in Q \times \mathbb{R}_+^{m+1} = K$, $u = (u_0, u_1, \ldots, u_m)^T$, и положим $f(z) = f_0(x) + u_0$, $\mathcal{F}(z) = (f_1(x) + u_1, \ldots, f_m(x) + u_m)^T$.

Рассмотрим задачу

$$f(z) \to \min, \quad \mathcal{F}(z) = 0, \quad z \in K.$$

Легко видеть, что если $\widehat{x}$ — глобальный минимум в задаче (P_5), то точка $\widehat{z} = (\widehat{x}, \widehat{u})$, где $\widehat{u} = (0, -f_1(\widehat{x}), \ldots, -f_m(\widehat{x}))^T$, является глобальным минимумом в этой задаче. Согласно предложению 5 §2.1 точка $\widehat{z}$ будет глобально экстремальной для отображения $F = (f, \mathcal{F})\colon K \to \mathbb{R}^{m+1}$.

Покажем, что отображение F является выпуклым, т. е. если $y_1, y_2 \in F(K)$ и $\alpha \in (0,1)$, то найдется такой элемент $z(\alpha) \in K$, что $F(z(\alpha)) = (1-\alpha)y_1 + \alpha y_2$.

Пусть $z_i = (x_i, u_i) \in K$, $i = 1, 2$, такие, что $y_i = F(z_i)$, $i = 1, 2$. Положим $G(x) = (f_0(x), f_1(x), \ldots, f_m(x))$, $z(\alpha) = (x(\alpha), u(\alpha))$, где $x(\alpha) = (1-\alpha)x_1 + \alpha x_2$ и $u(\alpha) = (1-\alpha)u_1 + \alpha u_2 + (1-\alpha)G(x_1) + \alpha G(x_2) - G(x(\alpha))$.

Заметим, что поскольку функции $f_0, f_1, \ldots, f_m$ выпуклы, то из неравенства Иенссена следует, что $(1-\alpha)G(x_1) + \alpha G(x_2) - G(x(\alpha)) \geq 0$ (неравенство покоординатное) и тем самым $u(\alpha) \in \mathbb{R}^{m+1}_+$.

Теперь имеем $(1-\alpha)y_1 + \alpha y_2 = (1-\alpha)F(x_1, u_1) + \alpha F(x_2, u_2) = (1-\alpha)G(x_1) + \alpha G(x_2) + \alpha u_1 + (1-\alpha)u_2 = G((1-\alpha)x_1 + \alpha x_2) + \alpha u_1 + (1-\alpha)u_2 + (1-\alpha)G(x_1) + \alpha G(x_2) - G((1-\alpha)x_1 + \alpha x_2) = G((1-\alpha)x_1 + \alpha x_2) + u(\alpha) = F(z(\alpha))$.

Итак, $\widehat{z}$ — глобально экстремальная точка для выпуклого отображения F. В силу предложения 7 найдется ненулевой вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ такой, что $\langle \lambda, F(z) - F(\widehat{z}) \rangle \geq 0$ для всех $z \in K$, или

$$\lambda_0(f_0(x) + u_0 - f_0(\widehat{x})) + \sum_{i=1}^{m} \lambda_i (f_i(x) + u_i) \geq 0 \tag{23}$$

для всех $(x, u) \in Q \times \mathbb{R}^{m+1}_+$, где $u = (u_0, u_1, \ldots u_m)$.

Полагая здесь $x = \widehat{x}$, $u_0 = 1$ и $u_i = -f_i(\widehat{x})$, $i = 1, \ldots, m$, получаем, что $\lambda_0 \geq 0$. Если $1 \leq j \leq m$, то полагая $x = \widehat{x}$, $u_0 = 0$, $u_i = -f_i(\widehat{x})$ при $i \neq j$ и $u_j = -f_j(\widehat{x}) + 1$, получаем, что $\lambda_j \geq 0$. Этим доказано утверждение (b) теоремы.

Так как (b) справедливо, то $\lambda_i f_i(\widehat{x}) \leq 0$, $i = 1, \ldots, m$. Но подставляя в (23) $x = \widehat{x}$ и $u = 0$, получаем, что $\sum_{i=1}^m \lambda_i f_i(\widehat{x}) \geq 0$ и поэтому $\lambda_i f_i(\widehat{x}) = 0$, $i = 1, \ldots, m$. Утверждение (c) доказано.

Пусть в (23) $u = 0$. Тогда, используя (c), будем иметь $\sum_{i=0}^m \lambda_i f_i(x) \geq \lambda_0 f_0(\widehat{x}) = \sum_{i=0}^m \lambda_i f_i(\widehat{x})$ для любого $x \in Q$. Это и есть утверждение (a) теоремы.

Докажем второе утверждение теоремы. Пусть x — допустимый вектор в задаче (P_5). Тогда, используя это обстоятельство вместе с

(b), затем (a) и (c), будем иметь

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

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)) $\mathcal{L}(\overline{x}, \lambda) = \sum_{i=1}^{m} \lambda_i f_i(\overline{x}) < 0 = \sum_{i=1}^{m} \lambda_i f_i(\widehat{x}) = \mathcal{L}(\widehat{x}, \lambda)$, что противоречит (a). $\square$

Глава 4. Условия экстремума в задачах вариационного исчисления и оптимального управления

§ 4.1. Задача Больца.

Необходимые условия первого порядка: уравнение Эйлера и условия трансверсальности.

Напомним, что через $C([t_0, t_1])$ и $C^1([t_0, t_1])$ обозначаются пространства непрерывных и непрерывно дифференцируемых функций $x(\cdot)$ на $[t_0, t_1]$. Это нормированные пространства соответственно с нормами $\|x(\cdot)\|_{C([t_0,t_1])} = \max_{t \in [t_0,t_1]} |x(t)|$ и $\|x(\cdot)\|_{C^1([t_0,t_1])} = \max(\|x(\cdot)\|_{C([t_0,t_1])}, \|\dot{x}(\cdot)\|_{C([t_0,t_1])})$.

Пусть $[t_0, t_1]$ — отрезок числовой прямой, G — открытое подмножество $\mathbb{R}^3$, W — открытое подмножество $\mathbb{R}^2$, $L\colon G \to \mathbb{R}$ — функция переменных t, x и $\dot{x}$ и $l\colon W \to \mathbb{R}$ — функция переменных ξ_0 и ξ_1. Задача

$$\mathcal{B}(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt + l(x(t_0), x(t_1)) \to \operatorname{extr} \qquad (P_6)$$

называется *задачей Больца.*

Функция $x(\cdot) \in C^1([t_0, t_1])$ называется *допустимой в задаче* (P_6), если $\Gamma(x(\cdot)) = \{(t, x(t), \dot{x}(t)) \mid t \in [t_0, t_1]\} \subset G$ и $(x(t_0), x(t_1) \in W$.

Допустимая функция $\widehat{x}(\cdot)$ называется *слабым локальным минимумом* (*максимумом*) в задаче (P_6), если существует такое $\varepsilon > 0$, что для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C^1([t_0,t_1])} < \varepsilon$ выполняется неравенство $\mathcal{B}(x(\cdot)) \geq \mathcal{B}(\widehat{x}(\cdot))$ $(\mathcal{B}(x(\cdot)) \leq \mathcal{B}(\widehat{x}(\cdot)))$.

Слабый локальный экстремум — это либо слабый локальный минимум, либо слабый локальный максимум.

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 $\widehat{l}_{\xi_i} = l_{\xi_i}(\widehat{x}(t_0), \widehat{x}(t_1))$, $i = 0, 1$.

Теорема (Необходимые условия экстремума в задаче (P_6))**.** *Пусть $\widehat{x}(\cdot)$ — локальный экстремум в задаче (P_6). Тогда, если функция L непрерывна вместе со своими частными производными по x и $\dot{x}$ на G, а функция l непрерывна вместе со своими частными производными по ξ_0 и ξ_1 на W, то $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0, t_1])$, 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 условие трансверсальности:

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

Доказательство. Пусть $h(\cdot) \in C^1([t_0, t_1])$. Положим $x(\alpha; h(\cdot))(\cdot) = \widehat{x}(\cdot) + \alpha h(\cdot)$, где $\alpha \in \mathbb{R}$. Ясно, что $x(\alpha; h(\cdot))(\cdot) \in C^1([t_0, t_1])$, что для достаточно малых по модулю α функция $x(\alpha; h(\cdot))(\cdot)$ допустима в задаче (P_6) и

$$\begin{aligned} \mathcal{B}(x(\alpha; h(\cdot))(\cdot)) = \int_{t_0}^{t_1} (L(t, \widehat{x}(t) + \alpha h(t), \dot{\widehat{x}}(t) + \alpha \dot{h}(t))\, dt \\ + l(\widehat{x}(t_0) + \alpha h(t_0), \widehat{x}(t_1) + \alpha h(t_1)). \end{aligned}$$

Очевидно, что функция $\alpha \mapsto \mathcal{B}(x(\alpha; h(\cdot))(\cdot))$ достигает в нуле локального экстремума. В силу условий теоремы, теоремы о производной суперпозиции функций и стандартных утверждений относительно дифференцируемости интеграла по параметру, эта функция дифференцируема в нуле и тогда по теореме Ферма ее производная в нуле равна нулю:

$$\begin{aligned} &\left.\frac{d\mathcal{B}(x(\alpha; h(\cdot))(\cdot))}{d\alpha}\right|_{\alpha=0} \\ &= \int_{t_0}^{t_1} \left(\widehat{L}_x(t)h(t) + \widehat{L}_{\dot{x}}(t)\dot{h}(t)\right) dt + \widehat{l}_{\xi_0} h(t_0) + \widehat{l}_{\xi_1} h(t_1) = 0. \end{aligned} \quad (24)$$

Интегрируя по частям первое слагаемое под интегралом, будем иметь

$$\int_{t_0}^{t_1} \widehat{L}_x(t)h(t)\,dt = \int_{t_0}^{t_1} h(t)d\left(-\int_t^{t_1} \widehat{L}_x(\tau)\,d\tau - \widehat{l}_{\xi_1}\right) = -\widehat{l}_{\xi_1}h(t_1)$$
$$+\left(\int_{t_0}^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1}\right)h(t_0) + \int_{t_0}^{t_1}\left(\int_t^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1}\right)\dot{h}(t)\,dt.$$

Подставляя это в предыдущее равенство, получим, что

$$\int_{t_0}^{t_1}\left(\int_t^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1} + \widehat{L}_{\dot{x}}(t)\right)\dot{h}(t)\,dt$$
$$+\left(\int_{t_0}^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1} + \widehat{l}_{\xi_0}\right)h(t_0) = 0.$$

Беря здесь функцию $\overline{h}(\cdot)$ такую, что $\dot{\overline{h}}(t) = \int_t^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1} + \widehat{L}_{\dot{x}}(t)$ и $\overline{h}(t_0) = \int_{t_0}^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1} + \widehat{l}_{\xi_0}$ (очевидно, $\overline{h}(\cdot) \in C^1([t_0,t_1])$), будем иметь

$$\int_{t_0}^{t_1}\left(\int_t^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1} + \widehat{L}_{\dot{x}}(t)\right)^2 dt$$
$$+\left(\int_{t_0}^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1} + \widehat{l}_{\xi_0}\right)^2 = 0.$$

Следовательно, $\int_t^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1} + \widehat{L}_{\dot{x}}(t) = 0$ для любого $t \in [t_0,t_1]$ и $\int_{t_0}^{t_1} \widehat{L}_x(\tau)\,d\tau + \widehat{l}_{\xi_1} + \widehat{l}_{\xi_0} = 0$. Из первого равенства вытекает, что $\widehat{L}_{\dot{x}}(\cdot) \in C^1([t_0,t_1])$, и дифференцируя его, получаем уравнение Эйлера. Далее, подставляя в него $t = t_1$, получаем, что $\widehat{L}_{\dot{x}}(t_1) = -\widehat{l}_{\xi_1}$, а подставляя $t = t_0$ и вычитая из второго равенства, получаем, что $\widehat{L}_{\dot{x}}(t_0) = \widehat{l}_{\xi_0}$. $\square$

Замечание. Наряду со слабым локальных экстремумом в задаче (P_6) можно рассматривать так называемый *сильный локальный экстремум.* Говорят, что допустимая в задаче (P_6) точка $\widehat{x}(\cdot)$ доставляет сильный локальный минимум (максимум), если существует такое $\varepsilon > 0$, что для любой допустимой функции $x(\cdot)$, для которой $\|x(\cdot) - \widehat{x}(\cdot)\|_{C([t_0,t_1])} < \varepsilon$ выполняется неравенство $\mathcal{B}(x(\cdot)) \geq \mathcal{B}(\widehat{x}(\cdot))$ ($\mathcal{B}(x(\cdot)) \leq \mathcal{B}(\widehat{x}(\cdot))$). При этом обычно рассматриваются функции не из пространства $C^1([t_0,t_1])$, а из более широкого пространства функций $PC^1([t_0,t_1])$, у которых производная лишь кусочно непрерывна. Фактически дословное повторение проведенных рассуждений показывает, что необходимыми условиями сильного локального экстремума также являются уравнение Эйлера и условия трансверсальности в задаче (P_6), рассматриваемой в пространстве $PC^1([t_0,t_1])$.

§ 4.2. Простейшая задача вариационного исчисления

4.2.1. Уравнение Эйлера — необходимое условие экстремума первого порядка.

Пусть $[t_0,t_1]$ — отрезок числовой прямой, G — открытое подмножество $\mathbb{R}^3$, $L\colon G\to\mathbb{R}$ — функция переменных t, x и $\dot x$ и $x_i\in\mathbb{R}$, $i=0,1$. Задача

$$J(x(\cdot))=\int_{t_0}^{t_1}L(t,x(t),\dot x(t))\,dt\to\mathrm{extr},\quad x(t_i)=x_i,\quad i=0,1\qquad (P_7)$$

называется *простейшей задачей* (*классического*) *вариационного исчисления.*

Эту задачу, как и задачу Больца, рассматриваем в пространстве $C^1([t_0,t_1])$. Понятия допустимой функции и слабого локального экстремума определяются аналогично тому как это было сделано для задачи Больца.

Теорема (Необходимые условия экстремума в задаче (P_7))**.** *Пусть $\widehat{x}(\cdot)$ — локальный экстремум в задаче (P_7). Тогда, если функция L непрерывна вместе со своими частными производными по x и $\dot x$ на G, то $\widehat{L}_{\dot x}(\cdot)\in C^1([t_0,t_1])$ и 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. Пусть $h(\cdot)\in C^1([t_0,t_1])$. Положим $x(\alpha;h(\cdot))(\cdot)=\widehat{x}(\cdot)+\alpha h(\cdot)$, где $\alpha\in\mathbb{R}$. Ясно, что $x(\alpha;h(\cdot))(\cdot)\in C^1([t_0,t_1])$ и найдется такое $r>0$, что $\Gamma(x(\alpha;h(\cdot))(\cdot))=\{(t,x(\alpha;h(\cdot))(t),\dot x(\alpha;h(\cdot))(t)) : t\in[t_0,t_1]\}\subset G$ для $\alpha\in B_{\mathbb{R}}(0,r)$. Рассмотрим задачу

$$\begin{gathered}
f_0(\alpha;h(\cdot))=J(x(\alpha;h(\cdot))(\cdot))\\
=\int_{t_0}^{t_1}L(t,\widehat{x}(t)+\alpha h(t),\dot{\widehat{x}}(t)+\alpha\dot h(t))\,dt\to\mathrm{extr},\\
f_1(\alpha;h(\cdot))=\widehat{x}(t_0)+\alpha h(t_0)-x_0=0,\\
f_2(\alpha;h(\cdot))=\widehat{x}(t_1)+\alpha h(t_1)-x_1=0,\quad \alpha\in B_{\mathbb{R}}(0,r).
\end{gathered}$$

Очевидно, что точка $\overline{\alpha}=0$ является локальным экстремумом в этой задаче. В силу предложения 5 из §2.1 эта точка является экстремальной для отображения $F\colon B_{\mathbb{R}}(0,r)\to\mathbb{R}^3$, $F(\alpha)=$

$F(\alpha; h(\cdot)) = (f_0(\alpha; h(\cdot)), f_1(\alpha; h(\cdot)), f_2(\alpha; h(\cdot)))$. Как было отмечено в §1.4 непрерывная дифференцируемость отображения влечет его строгую дифференцируемость, и поэтому по Основной теореме найдется ненулевой вектор $\lambda(h(\cdot)) = (\lambda_0(h(\cdot)), \lambda_1(h(\cdot)), \lambda_2(h(\cdot))) \in (\mathbb{R}^3)^*$ такой, что $\langle \lambda(h(\cdot)), F'(0; h(\cdot))\alpha \rangle \geq 0$ для любого $\alpha \in B_{\mathbb{R}}(0, r)$, или

$$(\lambda_0(h(\cdot))f_0'(0; h(\cdot)) + \lambda_1(h(\cdot))f_1'(0; h(\cdot)) + \lambda_2(h(\cdot))f_2'(0; h(\cdot)))\alpha \geq 0$$

для любого $\alpha \in B_{\mathbb{R}}(0, r)$. Отсюда, очевидно, следует, что множитель при α равен нулю, и тогда вычисляя производные, получаем

$$\lambda_0(h(\cdot)) \int_{t_0}^{t_1} \left(\widehat{L}_x(t)h(t) + \widehat{L}_{\dot{x}}(t)\dot{h}(t) \right) dt + \lambda_1(h(\cdot))h(t_0) \\ + \lambda_2(h(\cdot))h(t_1) = 0.$$

Это равенство справедливо для любого $h(\cdot) \in C^1([t_0, t_1])$. Оно является аналогом формулы (24) из доказательства предыдущей теоремы. Дословно повторяя рассуждения из этой теоремы, получаем, что найдется такое $\overline{h}(\cdot) \in C^1([t_0, t_1])$, для которого

$$\lambda_0(\overline{h}(\cdot)) \left(-\frac{d}{dt}\widehat{L}_{\dot{x}}(t) + \widehat{L}_x(t) \right) = 0, \quad \lambda_0(\overline{h}(\cdot))\widehat{L}_{\dot{x}}(t_0) = \lambda_1(\overline{h}(\cdot)), \\ \lambda_0(\overline{h}(\cdot))\widehat{L}_{\dot{x}}(t_1) = -\lambda_2(\overline{h}(\cdot)).$$

Отсюда следует, что $\lambda_0(\overline{h}(\cdot)) \neq 0$ (иначе вектор $\lambda(\overline{h}(\cdot))$ был бы нулевой, что невозможно) и тем самым уравнение Эйлера доказано. Заметим, что остальные соотношения не информативны и поэтому не участвуют в утверждении теоремы. □

Отметим, что здесь справедливо то же замечание, что и после предыдущей теоремы: уравнение Эйлера является необходимым условием сильного локального экстремума в задаче (P_7), рассматриваемой в пространстве $PC^1([t_0, t_1])$.

Мы рассмотрели "одномерные" варианты простейшей задачи и задачи Больца. Совершенно аналогично рассматриваются их векторные аналоги, когда $x = (x_1, \ldots, x_n)^T$. В этом случае роль пространств $C([t_0, t_1])$ и $C^1([t_0, t_1])$ играют пространства $C([t_0, t_1], \mathbb{R}^n)$ и $C^1([t_0, t_1], \mathbb{R}^n)$ — соответственно непрерывных и непрерывно дифференцируемых вектор-функций со значениями в $\mathbb{R}^n$. Они определяются аналогично одномерным вариантам, где $|x(t)| = \sqrt{x_1^2(t) + \ldots + x_n^2(t)}$. Необходимые условия экстремума здесь имеют тот же вид и их доказательства остаются прежними. Но формулы, разумеется, надо понимать векторно. Например,

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t)+\widehat{L}_x(t)=0 \ \Leftrightarrow \ -\frac{d}{dt}\widehat{L}_{\dot{x}_i}(t)+\widehat{L}_{x_i}(t)=0, \quad i=1,\ldots,n.$$

4.2.2. Необходимые условия экстремума второго порядка. *Квадратичные условия минимума. Условия минимума Лежандра и Якоби.*

Если фиксирована функция $\widehat{x}(\cdot)$, то для вторых частных производных (аналогично тому, как это было сделано для первых производных) используем обозначения: $\widehat{L}_{\dot{x}\dot{x}}(t) = L_{\dot{x}\dot{x}}(t,\widehat{x}(t),\dot{\widehat{x}}(t))$, $\widehat{L}_{\dot{x}x}(t)=L_{\dot{x}x}(t,\widehat{x}(t),\dot{\widehat{x}}(t))$ и $\widehat{L}_{xx}(t)=L_{xx}(t,\widehat{x}(t),\dot{\widehat{x}}(t))$.

Положим $C_0^1([t_0,t_1])=\{\,h(\cdot)\in C^1([t_0,t_1]) : h(t_0)=h(t_1)=0\,\}$.

Теорема (Квадратичные условия минимума для задачи (P_7))**.** *Пусть $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (P_7). Тогда, если функция L и ее частные производные по x и $\dot{x}$ до второго порядка включительно существуют и непрерывны на G, 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 $h(\cdot)\in C_0^1([t_0,t_1])$ справедливо неравенство

$$\int_{t_0}^{t_1}(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t)+2\widehat{L}_{\dot{x}x}(t)\dot{h}(t)h(t)+\widehat{L}_{xx}(t)h^2(t))\,dt\geq 0. \qquad (25)$$

Доказательство. Условия этой теоремы, очевидно, гарантируют выполнение условий предыдущей теоремы о необходимых условиях минимума первого порядка для задачи (P_7) и поэтому справедливо уравнение Эйлера.

Докажем второе утверждение теоремы. Пусть $h(\cdot)\in C_0^1([t_0,t_1])$. Положим $x(\alpha;h(\cdot))(\cdot)=\widehat{x}(\cdot)+\alpha h(\cdot)$, где $\alpha\in\mathbb{R}$. Ясно, что $x(\alpha;h(\cdot))(\cdot)\in C_0^1([t_0,t_1])$, $x(\alpha;h(\cdot))(t_i)=x_i$, $i=0,1$, и $\Gamma(x(\alpha;h(\cdot))(\cdot))=\{(t,x(\alpha;h(\cdot))(t),\dot{x}(\alpha;h(\cdot))(t)) : t\in[t_0,t_1]\}\subset G$ для достаточно малых по модулю α. Таким образом, для таких α функция $x(\alpha;h(\cdot))(\cdot)$ допустима в задаче (P_7).

В силу условий теоремы, теоремы о производной суперпозиции функций и стандартных утверждений относительно дифференцируемости интеграла по параметру, функция $\alpha\mapsto J(\alpha;h(\cdot))=$

$J(x(\alpha; h(\cdot))(\cdot))$ дифференцируема по α и ее производная имеет вид

$$J'(\alpha; h(\cdot)) = \int_{t_0}^{t_1} (L_{\dot{x}}(t,\ \widehat{x}(t) + \alpha h(t),\ \dot{\widehat{x}}(t) + \alpha \dot{h}(t))\dot{h}(t) + \\ + L_x(t,\ \widehat{x}(t) + \alpha h(t),\ \dot{\widehat{x}}(t) + \alpha \dot{h}(t))h(t))\, dt.$$

Как хорошо известно, необходимым условием минимума второго порядка для функции одного переменного $\alpha \mapsto J(\alpha; h(\cdot))$ является неотрицательность ее второй производной в нуле. Учитывая выражение для $J'(\alpha; h(\cdot))$, простые вычисления, возможные в силу предположений теоремы, показывают, что

$$J''(0; h(\cdot)) = \int_{t_0}^{t_1} \Big(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\widehat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) + \widehat{L}_{xx}(t)h^2(t)\Big)\, dt \geq 0.$$

Теорема доказана. □

Дальнейшее в этом пункте связано с проверкой неотрицательности выражения слева в (25) для всех функций $h(\cdot) \in C^1([t_0, t_1])$. Обозначим это выражение через $Q(\widehat{x}(\cdot))(h(\cdot))$.

Ясно, что условие $Q(\widehat{x}(\cdot))(h(\cdot)) \geq 0$ для всех $h(\cdot) \in C^1_0([t_0, t_1])$ равносильно тому, что функция $\widehat{h}(\cdot) = 0$ является минимумом задачи

$$Q(\widehat{x}(\cdot))(h(\cdot)) \to \min, \quad h(\cdot) \in C^1_0([t_0, t_1]). \tag{26}$$

В силу однородности функционала $h(\cdot) \mapsto Q(\widehat{x}(\cdot))(h(\cdot))$ любой его локальный минимум является и глобальным (проверьте это). Знакоопределенность этого функционала напрямую связана с вопросом существования нетривиального решения в задаче (26). Так как эта задача есть частный случай простейшей задачи вариационного исчисления, то для любого ее решения должно выполняться уравнение Эйлера, которое для данной задачи имеет вид

$$-\frac{d}{dt}\Big[\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}(t) + \widehat{L}_{\dot{x}x}(t)h(t)\Big] + \widehat{L}_{\dot{x}x}(t)\dot{h}(t) + \widehat{L}_{xx}(t)h(t) = 0. \tag{27}$$

Это дифференциальное уравнение называется *уравнением Якоби* для задачи (P_7).

Введем некоторые определения. Пусть $\widehat{x}(\cdot)$ — *экстремаль* задачи (P_7), т. е. $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера для этой задачи.

(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 $t \in [t_0, t_1]$.

(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 $h(t_0) = h(\tau) = 0$.

(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.

Теорема (Необходимое условие Лежандра)**.** *Пусть $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (P_7). Если интегрант L имеет непрерывные частные производные по x и $\dot{x}$ до второго порядка включительно на G, то выполнено условие Лежандра.*

Доказательство. Введем в пространстве $PC^1([t_0, t_1])$ (напомним, что это совокупность функций, у которых первая производная кусочно непрерывна) норму по формуле

$$\|h(\cdot)\| = \sqrt{|h(t_0)|^2 + \int_{t_0}^{t_1} \dot{h}^2(t)\, dt}.$$

Нам понадобится следующая

Лемма (о плотности)**.** *Множество $C_0^1([t_0, t_1])$ плотно в подпространстве $PC_0^1([t_0, t_1]) = \{\, h(\cdot) \in PC^1([t_0, t_1]) : h(t_0) = h(t_1) = 0 \,\}$.*

Доказательство. Покажем сначала, что $C^1([t_0, t_1])$ плотно в $PC^1([t_0, t_1])$. Действительно, пусть $h(\cdot) \in PC^1([t_0, t_1])$. Сопоставим $\dot{h}(\cdot)$ непрерывную функцию $x(\cdot)$, заменяя $\dot{h}(\cdot)$ в малых окрестностях ее точек разрыва наклонными прямыми. Тогда функция

$$\overline{h}(t) = h(t_0) + \int_{t_0}^{t} x(\tau)\, d\tau, \quad t \in [t_0, t_1]$$

принадлежит $C^1([t_0, t_1])$, и выбирая окрестности точек разрыва сколь угодно малыми, можно, очевидно, приблизить функцию $h(\cdot)$ в метрике $PC^1([t_0, t_1])$ с любой точностью функциями вида $\overline{h}(\cdot)$.

Обозначим для краткости $X = PC^1([t_0, t_1])$ и покажем, что множество $A_1 = \{\, h(\cdot) \in C^1([t_0, t_1]) : h(t_0) = 0 \,\}$ плотно в подпространстве $X_1 = \{\, h(\cdot) \in X : h(t_0) = 0 \,\}$ с индуцированной из X нормой. Пусть $h_0(\cdot) \in X_1$, $r > 0$ и $U_{X_1}(h_0(\cdot), r) = \{h(\cdot) \in X_1 : \|h(\cdot) - h_0(\cdot)\|_{X_1} < r\}$. По доказанному найдутся функции $h_i(\cdot) \in C^1([t_0, t_1])$, $i = 1, 2$, которые принадлежат соответственно открытым множествам $\{\, h(\cdot) \in X : h(t_0) < 0 \,\} \cap U_X(h_0(\cdot), r)$ и $\{\, h(\cdot) \in X : h(t_0) > 0 \,\} \cap U_X(h_0(\cdot), r)$. Ясно, что для любого $0 < \alpha < 1$ функция $h^\alpha(\cdot) = (1 - \alpha)h_1(\cdot) + \alpha h_2(\cdot)$ принадлежит $C^1([t_0, t_1])$ и содержится в $U_X(h_0(\cdot), r)$. Так как $h^0(t_0) < 0$, а $h^1(t_0) > 0$, то $h^{\alpha_0}(t_0) = 0$ для некоторого $\alpha_0 \in (0,1)$, т. е. функция $h^{\alpha_0}(\cdot)$ принадлежит A и содержится в шаре $U_{X_1}(h_0(\cdot), r)$. Это означает, в силу произвольности $h_0(\cdot)$ и r, что A плотно в X_1.

Если теперь взять за исходное пространство X_1, рассмотреть в нем подпространство $X_2 = \{h(\cdot) \in X_1 : h(t_1) = 0\}$ и рассмотреть множество $A_2 = \{h(\cdot) \in A_1 : h(t_1) = 0\}$, то рассуждая точно также как и выше, получим, что A_2 плотно в X_2, а это то, что и требовалось доказать. $\square$

Из квадратичных условий минимума следует, что $Q(\widehat{x}(\cdot))(h(\cdot)) \geq 0$ для любого $h(\cdot) \in C_0^1([t_0, t_1])$. Очевидно, что величина $Q(\widehat{x}(\cdot))(h(\cdot))$ имеет смысл для любого $h(\cdot) \in PC^1([t_0, t_1])$ и нетрудно проверить, что функционал $h(\cdot) \mapsto Q(\widehat{x}(\cdot))(h(\cdot))$ непрерывен на $PC^1([t_0, t_1])$. Тогда из леммы получаем, что $Q(\widehat{x}(\cdot))(h(\cdot)) \geq 0$ для всех $h(\cdot) \in PC_0^1([t_0, t_1])$.

Теорему доказываем от противного. Пусть существует такая точка $\tau \in [t_0, t_1]$, 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 $v > 0$ 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 $t \in [\tau - \varepsilon_0/2, \tau + \varepsilon_0/2]$. 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 $\widetilde{h}_\varepsilon(\cdot) \in PC_0^1([t_0, t_1])$ и

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

Оценим слагаемые в выражении для $Q(\widehat{x}(\cdot))(\widetilde{h}_\varepsilon(\cdot))$. Имеем

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

Далее, так как $\max_{t \in [t_0, t_1]} \widetilde{h}_\varepsilon(t) = v\varepsilon/2$ и $|\dot{\widetilde{h}}_\varepsilon(t)| \leq v$, то

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

Наконец,

$$\int_{t_0}^{t_1} \widehat{L}_{xx}(t)\widetilde{h}_\varepsilon^2(t)\,dt \le \frac{v^2\varepsilon^2}{4}\int_{t_0}^{t_1} |\widehat{L}_{xx}(t)|\,dt \le C_1\varepsilon^2,$$

где $C_1 = (v^2/4)(t_1-t_0)\max_{t\in[t_0,t_1]}|\widehat{L}_{xx}(t)|$.

Учитывая полученные оценки, непосредственный подсчет показывает, что $Q(\widehat{x}(\cdot))(\widetilde{h}_\varepsilon(\cdot)) < -\varepsilon+2C\varepsilon^2+C_1\varepsilon^2$, т. е. $Q(\widehat{x}(\cdot))(\widetilde{h}_\varepsilon(\cdot)) < 0$ для достаточно малого ε в противоречие с тем, что $Q(\widehat{x}(\cdot))(h(\cdot)) \ge 0$ для всех $h(\cdot) \in PC_0^1([t_0,t_1])$. $\square$

Для дальнейших рассмотрений понадобится следующая конструкция. Пусть $x(\cdot)$ — решение уравнения Якоби (27). Положим $y(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)\dot{x}(\cdot) + \widehat{L}_{\dot{x}x}(\cdot)x(\cdot)$. Согласно (27) это дифференцируемая функция и $\dot{y}(\cdot) = \widehat{L}_{\dot{x}x}(\cdot)\dot{x}(\cdot)+\widehat{L}_{xx}(\cdot)x(\cdot)$. Пусть выполнено усиленное условие Лежандра. Тогда выражая $\dot{x}(\cdot)$ из первого соотношения, а затем подставляя его во второе, получаем, что пара $(x(\cdot), y(\cdot))$ является решение следующей системы линейных уравнений

$$\begin{cases} \dot{x} = -\dfrac{\widehat{L}_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}\,x + \dfrac{1}{\widehat{L}_{\dot{x}\dot{x}}(t)}\,y \\[2ex] \dot{y} = \left(\widehat{L}_{xx}(t) - \dfrac{\widehat{L}^2_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}\right)x + \dfrac{\widehat{L}_{\dot{x}x}(t)}{\widehat{L}_{\dot{x}\dot{x}}(t)}\,y. \end{cases} \tag{28}$$

Верно, очевидно, и обратное, если пара $(x(\cdot), y(\cdot))$ — решение системы (28), то $x(\cdot)$ — решение уравнения Якоби (27).

Как хорошо известно из теории обыкновенных дифференциальных уравнений, система (28) для любых начальных условий имеет единственное решение, определенное на на всем отрезке $[t_0,t_1]$.

Теорема (Необходимое условие Якоби). *Пусть $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (P_7). Если интегрант L имеет непрерывные частные производные по x и $\dot{x}$ до второго порядка включительно на G и выполнено усиленное условие Лежандра, то выполнено условие Якоби.*

Доказательство. Доказываем от противного. Пусть $\tau \in (t_0,t_1)$ и $\overline{h}(\cdot) \in C_0^1([t_0,t_1])$ — нетривиальное решение уравнения Якоби такое, что $\overline{h}(\tau) = 0$. Положим

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

Интегрируя по частям, будем иметь

$$Q(\widehat{x}(\cdot))(\overline{h}(\cdot)) = \int_{t_0}^{\tau} \left(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}^2(t) + 2\widehat{L}_{\dot{x}x}(t)\dot{\overline{h}}(t)\overline{h}(t) + \widehat{L}_{xx}(t)\overline{h}^2(t) \right) dt$$

$$= \int_{t_0}^{\tau} \left(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}(t) + \widehat{L}_{\dot{x}x}(t)\overline{h}(t) \right) \dot{\overline{h}}(t)\, dt + \int_{t_0}^{\tau} \left(\widehat{L}_{x\dot{x}}(t)\dot{\overline{h}}(t) \right.$$

$$\left. + \widehat{L}_{xx}(t)\overline{h}(t) \right) \overline{h}(t)\, dt = \int_{t_0}^{\tau} \left(-\frac{d}{dt} \left(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{\overline{h}}(t) \right.\right.$$

$$\left.\left. + \widehat{L}_{\dot{x}x}(t)\overline{h}(t) \right) + \widehat{L}_{x\dot{x}}(t)\dot{\overline{h}}(t) + \widehat{L}_{xx}(t)\overline{h}(t) \right) \overline{h}(t)\, dt.$$

Поскольку $\overline{h}(\cdot)$ удовлетворяет уравнению Якоби, то отсюда следует, что $Q(\widehat{x}(\cdot))(\widetilde{h}(\cdot)) = 0$. В предыдущей теореме показано, что $Q(\widehat{x}(\cdot))(h(\cdot)) \geq 0$ для всех $h(\cdot) \in PC_0^1([t_0, t_1])$ (см. рассуждения после доказательства леммы о плотности). Поэтому функция $\widetilde{h}(\cdot)$ является решением задачи (26), но для функций из $PC_0^1([t_0, t_1])$. В этом случае (см. замечание после доказательства теоремы о необходимых условиях в задаче (P_7)) эта функция также удовлетворяет уравнению Эйлера, т. е. уравнению (27). Покажем, что 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, $\widetilde{h}(\cdot) = \dot{\widetilde{h}}(\cdot) = 0$ 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 пара $(x(\cdot), y(\cdot))$, где $x(\cdot) = \overline{h}(\cdot)$ и $y(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)\dot{\overline{h}}(\cdot) + \widehat{L}_{\dot{x}x}(\cdot)\overline{h}(\cdot)$, была решение системы (28) с начальными условиями $x(\tau) = y(\tau) = 0$ и значит, в силу единственности $x(\cdot) = \overline{h}(\cdot) = 0$, 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. $\square$

4.2.3. Достаточные условия экстремума второго порядка.

Достаточные условия минимума второго порядка в терминах квадратичной формы и в форме усиленных условий Лежандра и Якоби.

Теорема (Достаточные условия слабого минимума в задаче (P_7) в терминах квадратичной формы)**.** *Пусть интегрант L дважды непрерывно дифференцируем на открытом множестве $G \subset \mathbb{R}^3$*

и $\widehat{x}(\cdot)$ — допустимая функция в задаче (P_7). Тогда, если $\widehat{x}(\cdot)$ для каждого $t \in [t_0, t_1]$ удовлетворяет уравнению Эйлера

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

и существует такое $\alpha > 0$, что для всех $h(\cdot) \in C_0^1([t_0, t_1])$ справедливо неравенство

$$\int_{t_0}^{t_1} \Big(\widehat{L}_{\dot{x}\dot{x}}(t)\dot{h}^2(t) + 2\widehat{L}_{\dot{x}x}(t)\dot{h}(t)h(t) + \widehat{L}_{xx}(t)h^2(t)\Big)\, dt \geq \alpha \int_{t_0}^{t_1} \dot{h}^2(t)\, dt, \quad (29)$$

то $\widehat{x}(\cdot)$ доставляет слабый локальный минимум в задаче (P_7).

Доказательство. Покажем сначала, что существует такая константа $c > 0$, что для любой функции $h(\cdot) \in C^1([t_0, t_1])$, для которой $h(t_0) = 0$, справедливо неравенство

$$\frac{1}{2}\int_{t_0}^{t_1} (\dot{h}^2(t) + 2|\dot{h}(t)h(t)| + h^2(t))\, dt \leq c\int_{t_0}^{t_1} \dot{h}^2(t)\, dt. \qquad (30)$$

Действительно, используя неравенство Коши–Буняковского, будем иметь для любого $t \in [t_0, t_1]$

$$|h(t)| = \left|\int_{t_0}^{t} \dot{h}(\tau)\, d\tau\right| \leq \int_{t_0}^{t_1} |\dot{h}(t)|\, dt \leq \sqrt{t_1 - t_0}\left(\int_{t_0}^{t_1} \dot{h}^2(t)\, dt\right)^{1/2}.$$

Следовательно,

$$\int_{t_0}^{t_1} h^2(t)\, dt \leq (t_1 - t_0)\max_{t\in[t_0,t_1]} |h(t)|^2 \leq (t_1 - t_0)^2\int_{t_0}^{t_1} \dot{h}^2(t)\, dt.$$

Так как, очевидно, $|\dot{h}(t)h(t)| \leq (\dot{h}^2(t) + h^2(t))/2$, то

$$\frac{1}{2}\int_{t_0}^{t_1} (\dot{h}^2(t) + 2|\dot{h}(t)h(t)| + h^2(t))\, dt \leq \int_{t_0}^{t_1} (\dot{h}^2(t) + h^2(t))\, dt \leq (1 + (t_1 - t_0)^2)\int_{t_0}^{t_1} \dot{h}^2(t)\, dt = c\int_{t_0}^{t_1} \dot{h}^2(t)\, dt.$$

Пусть теперь $\varepsilon > 0$ таково, что $\varepsilon c \leq \alpha$. Нетрудно проверить, что существует такое $\delta_0 > 0$, что компакт $K = \{\, (t, x, \dot{x}) : |x - \widehat{x}(t)| + |\dot{x} - \dot{\widehat{x}}(t)| \leq \delta_0,\ t \in [t_0, t_1]\,\}$ принадлежит G. Функции $L_{\dot{x}\dot{x}}(t, x, \dot{x})$,

$L_{\dot{x}x}(t,x,\dot{x})$ и $L_{xx}(t,x,\dot{x})$ равномерно непрерывны на K. Следовательно найдется $0<\delta=\delta(\varepsilon)\le\delta_0$ такое, что

$$|L_{\dot{x}\dot{x}}(t,x_1,\dot{x}_1)-L_{\dot{x}\dot{x}}(t,x_2,\dot{x}_2)|<\varepsilon,\ \ |L_{\dot{x}x}(t,x_1,\dot{x}_1)-L_{\dot{x}x}(t,x_2,\dot{x}_2)|<\varepsilon,$$
$$|L_{xx}(t,x_1,\dot{x}_1)-L_{xx}(t,x_2,\dot{x}_2)|<\varepsilon \quad (31)$$

для $(t,x_i,\dot{x}_i)\in K$, $i=1,2$, для которых $|x_1-x_2|<\delta$ и $|\dot{x}_1-\dot{x}_2|<\delta$.

Пусть функция $x(\cdot)$ допустима в задаче (P_7) и $\|x(\cdot)-\widehat{x}(\cdot)\|_{C^1([t_0,t_1])}<\delta$. Положим $h(\cdot)=x(\cdot)-\widehat{x}(\cdot)$ (тогда $h(\cdot)\in C_0^1([t_0,t_1])$). Для всех $t\in[t_0,t_1]$ точки $(t,\widehat{x}(t)+h(t),\dot{\widehat{x}}(t)+\dot{h}(t))$ принадлежат K и тем самым G. По формуле Тейлора

$$\begin{aligned}L(t,\widehat{x}(t)+h(t),\dot{\widehat{x}}(t)+\dot{h}(t))&=\widehat{L}(t)+\widehat{L}_x(t)h(t)+\widehat{L}_{\dot{x}}(t)\dot{h}(t)\\&+\frac{1}{2}(L_{xx}(t,\widehat{x}(t)+\theta h(t),\dot{\widehat{x}}(t)+\theta\dot{h}(t))h^2(t)\\&+2L_{x\dot{x}}(t,\widehat{x}(t)+\theta h(t),\dot{\widehat{x}}(t)+\theta\dot{h}(t))h(t)\dot{h}(t)\\&+L_{\dot{x}\dot{x}}(t,\widehat{x}(t)+\theta h(t),\dot{\widehat{x}}(t)+\theta\dot{h}(t))\dot{h}^2(t)),\end{aligned}\quad (32)$$

где $\theta=\theta(t)\in(0,1)$.

Обозначим через $q(t)$ выражение под знаком интеграла слева в (29). Прибавим и вычтем величину $(1/2)q(t)$ из правой части равенства (32). Тогда в силу соотношений (31), будем иметь

$$\begin{aligned}L(t,\widehat{x}(t)+h(t),\dot{\widehat{x}}(t)+\dot{h}(t))-\widehat{L}(t)&\ge\widehat{L}_x(t)h(t)+\widehat{L}_{\dot{x}}(t)\dot{h}(t)\\&+\frac{1}{2}q(t)-\frac{\varepsilon}{2}(h^2(t)+2|h(t)\dot{h}(t)|+\dot{h}^2(t)).\end{aligned}$$

Интегрируя это соотношение, затем интегрируя первое слагаемое справа по частям и используя оценки (29) и (30), получим

$$\begin{aligned}J(x(\cdot))-J(\widehat{x}(\cdot))&\ge\int_{t_0}^{t_1}\left(-\int_{t_0}^{t}\widehat{L}_x(\tau)\,d\tau+\widehat{L}_{\dot{x}}(t)\right)\dot{h}(t)\,dt\\&+\frac{\alpha}{2}\int_{t_0}^{t_1}\dot{h}^2(t)\,dt-\frac{\varepsilon c}{2}\int_{t_0}^{t_1}\dot{h}^2(t)\,dt\end{aligned}$$

Первое слагаемое справа здесь равно нулю в силу того, что $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера. Сумма остальных слагаемых неотрицательна в силу выбора ε. Таким образом, $J(x(\cdot))\ge J(\widehat{x}(\cdot))$ для всех допустимых $x(\cdot)$ таких, что $\|x(\cdot)-\widehat{x}(\cdot)\|_{C^1([t_0,t_1])}<\delta$ и значит, $\widehat{x}(\cdot)$ — слабый локальный минимум в задаче (P_7). $\square$

Теорема (Достаточные условия слабого минимума в простейшей задаче в терминах условий Лежандра и Якоби)**.** *Пусть интегрант L непрерывен вместе со своими частными производными* *до третьего порядка включительно на открытом множестве $G \subset \mathbb{R}^3$, допустимая в задаче (P_7) функция $\widehat{x}(\cdot)$ такова, что $\widehat{x}(\cdot) \in C^2([t_0, t_1])$, $\widehat{x}(\cdot)$ удовлетворяет уравнению Эйлера и на $\widehat{x}(\cdot)$ выполнены усиленные условия Лежандра и Якоби. Тогда $\widehat{x}(\cdot)$ доставляет слабый локальный минимум в задаче (P_7).*

Доказательство. Покажем, что в условиях теоремы справедливо неравенство (29). Тогда из теоремы о достаточных условиях в терминах квадратичной формы, очевидно, будет следовать утверждение данной теоремы.

Пусть $(x(\cdot), y(\cdot))$ — решение (28) с условиями: $x(t_0) = 0$, $y(t_0) = 1$. Ясно, что $x(\cdot) \neq 0$, $y(\cdot) \neq 0$ и из первого уравнения следует, что $\dot{x}(t_0) = \widehat{L}^{-1}_{\dot{x}\dot{x}}(t_0) > 0$. Как уже было отмечено ранее, $x(\cdot)$ является решением уравнения Якоби (см. (27)).

Обозначим для краткости $A(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)$, $B(\cdot) = \widehat{L}_{\dot{x}x}(\cdot)$, $C(\cdot) = \widehat{L}_{xx}(\cdot)$ и покажем, что

$$Q(h(\cdot)) = \int_{t_0}^{t_1} \Big(A(t)\dot{h}^2(t) + 2B(t)\dot{h}(t)h(t) + C(t)h^2(t)\Big)\, dt \geq 0$$

для всех $h(\cdot) \in C_0^1([t_0, t_1])$. Для этого покажем, что справедлива формула

$$Q(h(\cdot)) = \int_{t_0}^{t_1} A(t) \left(\dot{h}(t) - \frac{\dot{x}(t)}{x(t)}\, h(t)\right)^2 dt. \tag{33}$$

Заметим, что величина справа имеет смысл, поскольку в силу усиленного условия Якоби функция $x(\cdot)$ не обращается в ноль на $(t_0, t_1]$, а отношение $h(t)/x(t)$ стремится к конечному пределу при $t \to t_0$ (согласно правилу Лопиталя).

Обозначим через $I(h(\cdot))$ интеграл справа в (33). Имеем

$$I(h(\cdot)) = \int_{t_0}^{t_1} \left(A(t)\dot{h}^2(t) - 2A(t)\frac{\dot{x}(t)}{x(t)}\dot{h}(t)h(t) + A(t)\frac{\dot{x}^2(t)}{x^2(t)}h^2(t)\right) dt.$$

Проинтегрируем по частям во втором слагаемом. Это возможно, так как в уравнении Якоби, которое в новых обозначениях имеет вид

$$-\frac{d}{dt}\Big[A(t)(t)\dot{h}(t) + B(t)h(t)\Big] + B(t)\dot{h}(t) + C(t)h(t) = 0, \tag{34}$$

выражение в квадратных скобках дифференцируемо, функция $B(\cdot)x(\cdot)$ дифференцируема в силу условий теоремы, и поэтому функция $A(\cdot)\dot{x}(\cdot)$ дифференцируема, а значит и функция

$A(\cdot)\dot{x}(\cdot)/x(\cdot)$ дифференцируема на $(t_0, t_1]$. Итак, имеем после интегрирования по частям

$$I(h(\cdot)) = \int_{t_0}^{t_1} \left(A(t)\dot{h}^2(t) + \left(\frac{d}{dt}\left(A(t)\frac{\dot{x}(t)}{x(t)} \right) + A(t)\frac{\dot{x}^2(t)}{x^2(t)} \right) h^2(t) \right) dt.$$

Преобразуем выражение перед $h^2(t)$, дифференцируя первое слагаемое, как отношение двух функций, затем используя уравнение (34) и формулу дифференцирования произведения функций, будем иметь (штрих, как и точка, обозначает производную по t)

$$\frac{(A(t)\dot{x}(t))'x(t) - \dot{x}(t)A(t)\dot{x}(t)}{x^2(t)} + A(t)\frac{\dot{x}^2(t)}{x^2(t)} = \frac{(A(t)\dot{x}(t))'}{x(t)}$$
$$= \frac{-(B(t)x(t))' + B(t)\dot{x}(t) + C(t)x(t)}{x(t)} = -\dot{B}(t) + C(t).$$

Подставляя полученное в предыдущее равенство и интегрируя по частям, приходим к соотношению

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

которое означает, что выполнено равенство (33) и тем самым $Q(h(\cdot)) \geq 0$ на $C_0^1([t_0, t_1])$.

Покажем теперь, что отсюда следует (29).

Для любого $0 < \alpha < \min_{t\in[t_0,t_1]} A(t)$ функция $A_\alpha(\cdot) = A(\cdot) - \alpha$ положительна на $[t_0, t_1]$. Для таких α рассмотрим систему (28) с заменой $A(\cdot) = \widehat{L}_{\dot{x}\dot{x}}(\cdot)$ на $A_\alpha(\cdot)$. Пусть $(x_\alpha(\cdot), y_\alpha(\cdot))$ — ее решение с начальными условиями: $x_\alpha(t_0) = 0$, $y_\alpha(t_0) = 1$. Из первого уравнения этой системы и так как $A_\alpha(t_0) < A(t_0)$ получаем, что $\dot{x}_\alpha(t_0) = A_\alpha^{-1}(t_0) > A^{-1}(t_0) > 0$. Отсюда легко следует, что существует такое $\varepsilon > 0$, что $x_\alpha(t) > 0$ для $t \in (t_0, \varepsilon]$ и всех указанных α. По теореме о непрерывной зависимости решения дифференциального уравнения от параметров вытекает, что $x_\alpha(\cdot) \to x(\cdot)$ при $\alpha \to 0$ в $C([t_0, t_1])$ и тогда для достаточно малых $\alpha > 0$ функции $x_\alpha(\cdot)$ будут положительны на $[\varepsilon, , t_1]$ и значит, на $(t_0, t_1]$. Фиксируем такое α и проводя те же рассуждения, что и выше с заменой $x(\cdot)$ на $x_\alpha(\cdot)$, получаем, что

$$Q_\alpha(h(\cdot)) = \int_{t_0}^{t_1} (A_\alpha(t)\dot{h}^2(t) + 2B(t)\dot{h}(t)h(t) + C(t)h^2(t))\, dt \geq 0$$

для всех $h(\cdot) \in C_0^1([t_0, t_1])$. Следовательно,

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

т. е. справедливо неравенство (29), что, как было отмечено выше, доказывает теорему. $\square$

§ 4.3. Задача Лагранжа вариационного исчисления.

4.3.1. Постановка задачи.

Пусть $[t_0, t_1]$ — конечный отрезок, G — открытое подмножество $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r$, W — открытое подмножество $\mathbb{R}^n \times \mathbb{R}^n$, заданы функции $f_i \colon G \to \mathbb{R}$, $i = 0, 1, \ldots, m$, и отображение $\varphi \colon G \to \mathbb{R}^n$ (переменных $t \in \mathbb{R}$, $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $u = (u_1, \ldots, u_r)^T \in \mathbb{R}^r$) и функции $l_i \colon W \to \mathbb{R}$, $i = 0, 1, \ldots, m$ переменных ξ_0 и ξ_1. Обозначим

$$\mathcal{B}_i(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} f_i(t, x(t), u(t))\, dt + l_i(x(t_0), x(t_1)), \quad i = 0, 1, \ldots, m.$$

В пространстве $Z = C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^r)$,[4] относительно переменной $(x(\cdot), u(\cdot)) \in Z$ рассмотрим задачу

$$\begin{aligned} \mathcal{B}_0(x(\cdot), u(\cdot)) \to \min, \quad \dot{x}(t) = \varphi(t, x(t), u(t)), \\ \mathcal{B}_i(x(\cdot), u(\cdot)) \leq 0, \quad 1 \leq i \leq m', \\ \mathcal{B}_i(x(\cdot), u(\cdot)) = 0, \quad m' + 1 \leq i \leq m, \end{aligned} \qquad (P_8)$$

которая называется *задачей Лагранжа вариационного исчисления* (в понтрягинской форме).

Пара $(x(\cdot), u(\cdot))$ называется *допустимой в задаче* (P_8), если $\{\, (t, x(t), u(t)) \,:\, t \in [t_0, t_1] \,\} \subset G$, $(x(t_0), x(t_1)) \in W$ и $(x(\cdot), u(\cdot))$ удовлетворяет ограничениям задачи.

Допустимая пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ называется *слабым локальным минимумом в задаче* (P_8), если существует такая окрестность $\mathcal{O}$ пары $(x(\cdot), u(\cdot))$, что для любой допустимой пары $(x(\cdot), u(\cdot)) \in \mathcal{O}$ выполнено неравенство $\mathcal{B}_0(x(\cdot), u(\cdot)) \geq \mathcal{B}_0(\widehat{x}(\cdot), \widehat{u}(\cdot))$.

4.3.2. Уравнения Эйлера-Лагранжа - необходимые условия минимума первого порядка.

[4] Z — нормированное пространство с нормой $\|(x(\cdot), u(\cdot))\|_Z = \|x(\cdot)\|_{C^1([t_0,t_1],\mathbb{R}^n)} + \|u(\cdot)\|_{C([t_0,t_1],\mathbb{R}^n)}$.

Положим $L(t,x,\dot{x},u,\lambda,p)=\sum_{i=0}^{m}\lambda_i f_i(t,x,u)+\langle p,\dot{x}-\varphi(t,x,u)\rangle$, $l(\xi_0,\xi_1)=\sum_{i=0}^{m}\lambda_i l_i(\xi_0,\xi_1)$, где $\lambda=(\lambda_0,\lambda_1,\ldots,\lambda_m)\in(\mathbb{R}^{m+1})^*$ и $p\in(\mathbb{R}^n)^*$.

Если фиксирована пара $(\widehat{x}(t),\widehat{u}(t))$, то для сокращения записи пишем $\widehat{\varphi}(t)=\varphi(t,\widehat{x}(t),\widehat{u}(t))$, $\widehat{f_i}(t)=f_i(t,\widehat{x}(t),\widehat{u}(t))$, $i=0,1,\ldots,m$, и аналогично для производных по x и u этих функций, $\widehat{l}_{\zeta_i}=l_{\zeta_i}(\widehat{x}(t_0),\widehat{x}(t_1))$, $i=0,1$. Если еще фиксирована функция $p=p(\cdot)$ на $[t_0,t_1]$ со значениями в $(\mathbb{R}^n)^*$, то пишем $\widehat{L}(t,p)=\sum_{i=0}^{m}\lambda_i\widehat{f_i}(t)+\langle p(t),\dot{\widehat{x}}(t)-\widehat{\varphi}(t)\rangle$ и аналогично для производных по x, $\dot{x}$ и u этой функции.

Теорема (Необходимые условия минимума в задаче Лагранжа (P_8))**.** *Пусть $(\widehat{x}(\cdot),\widehat{u}(\cdot))$ доставляет слабый локальный минимум в задаче (P_8). Тогда, если функции f_i, $0\le i\le m$, и отображение φ непрерывны вместе со своими частными производными по x и u на G, а функции l_i, $0\le i\le m$, непрерывно дифференцируемы на W, то найдутся вектор $\lambda=(\lambda_0,\lambda_1,\ldots,\lambda_m)\in(\mathbb{R}^{m+1})^*$ и функция $p(\cdot)\in C^1([t_0,t_1],(\mathbb{R}^n)^*)$, не равные одновременно нулю, такие, что $\widehat{L}_{\dot{x}}(\cdot)\in C^1([t_0,t_1],\mathbb{R}^n)$ и выполняются*

(a) *уравнение Эйлера–Лагранжа (условие стационарности по x)*

$$-\frac{d}{dt}\widehat{L}_{\dot{x}}(t,p)+\widehat{L}_x(t,p)=0 \;\;\Leftrightarrow\;\; -\dot{p}(t)=p(t)\widehat{\varphi}_x(t)-\sum_{i=0}^{m}\lambda_i\widehat{f}_{ix}(t),$$

(b) *условия трансверсальности*

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

(c) *условие стационарности по u*

$$\widehat{L}_u(t,p)=0 \;\;\Leftrightarrow\;\; p(t)\widehat{\varphi}_u(t)-\sum_{i=0}^{m}\lambda_i\widehat{f}_{iu}(t)=0,$$

(d) *условия дополняющей нежесткости*

$$\lambda_i\mathcal{B}_i(\widehat{x}(\cdot),\widehat{u}(\cdot))=0,\quad i=1,\ldots,m',$$

(e) *условия неотрицательности*

$$\lambda_i\ge 0,\quad i=0,1,\ldots m'.$$

Чтобы не утяжелять изложение, доказательство проведем для частного случая задачи (P_8), а именно, рассмотрим такую постановку

$$\int_{t_0}^{t_1} f_0(t, x(t), u(t))\, dt \to \min, \quad \dot{x} = \varphi(t, x, u),$$
$$x(t_0) = x_0, \quad g_i(x(t_1)) = 0, \quad i = 1, \ldots, m, \quad (P_8')$$

где g_i, $i = 1, \ldots, m$, — непрерывно дифференцируемые функции на некотором открытом множестве в $\mathbb{R}^n$.

В данном случае $L(t, x, \dot{x}, u, \lambda_0, p) = \lambda_0 f_0(t, x, u) + \langle p, \dot{x} - \varphi(t, x, u)\rangle$, $l(x(t_0), x(t_1)) = \langle \mu, x(t_0)\rangle + \sum_{i=1}^m \lambda_i g_i(x(t_1))$, где $\mu \in (\mathbb{R}^n)^*$, $\overline{\lambda} = (\lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^m)^*$ и $p \in (\mathbb{R}^n)^*$.

Тогда утверждения теоремы для задачи (P_8'), как легко проверить, будут таковы: утверждения (a) и (c) останутся прежними, но с $m = 0$, условия дополняющей нежесткости отсутствуют, $\lambda_0 \geq 0$, а условия трансверсальности имеют вид

$$\widehat{L}_{\dot{x}}(t_1, p) = -\sum_{i=1}^m \lambda_i g_i'(\widehat{x}(t_1)) \;\; \Leftrightarrow \;\; p(t_1) = -\sum_{i=1}^m \lambda_i g_i'(\widehat{x}(t_1)).$$

По поводу последнего заметим, что формально следовало бы добавить еще соотношение $\widehat{L}_{\dot{x}}(t_0, p) = p(t_0) = \mu$, но это условие мы не пишем, так как оно не несет никакой информации, поскольку μ ни в какие другие соотношения не входит и если $\lambda_0 = 0$, $\overline{\lambda} = 0$ и $p = 0$, то $\mu = 0$.

Перед непосредственным выводом необходимых условий минимума в задаче (P_8') приведем эту задачу к задаче без интегрального функционала, исследование которой технически более удобно. Для этого обозначим $\varphi_0 = f_0$, $\overline{\varphi} = (\varphi_0, \varphi)^T = (\varphi_0, \varphi_1, \ldots, \varphi_n)^T$ и $\overline{x} = (x_0, x)^T = (x_0, x_1, \ldots, x_n)^T$.

Рассмотрим задачу

$$x_0(t_1) \to \min, \quad \dot{\overline{x}}(t) = \overline{\varphi}(t, \overline{x}(t), u(t)),$$
$$x_0(t_0) = 0, \quad x(t_0) = x_0, \quad g_i(x(t_1)) = 0, \quad i = 1, \ldots, m. \quad (35)$$

Заметим, что если $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — слабый локальный минимум в задаче (P_8'), то $(\widehat{\overline{x}}(\cdot), \widehat{u}(\cdot)) = ((\widehat{x}_0(\cdot), \widehat{x}(\cdot)), \widehat{u}(\cdot))$, где $\widehat{x}_0(t) = \int_{t_0}^t f_0(\tau, \widehat{x}(\tau), \widehat{u}(\tau))\, d\tau$, — слабый локальный минимум в задаче (35).

Действительно, пусть $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — слабый локальный минимум в (P_8') и окрестность $\mathcal{O}$ точки $(\widehat{x}(\cdot), \widehat{u}(\cdot)) \in Z$ такова, что для любой допустимой пары $(x(\cdot), u(\cdot)) \in \mathcal{O}$ выполнено неравенство $\int_{t_0}^{t_1} f_0(t, x(t), u(t))\, dt \geq \int_{t_0}^{t_1} f_0(t, \widehat{x}(t), \widehat{u}(t))\, dt$.

Рассмотрим окрестность $\mathcal{O}_1 = C^1([t_0,t_1]) \times \mathcal{O}$ пары $(\widehat{\overline{x}}(\cdot), \widehat{u}(\cdot))$. Если пара $(\overline{x}(\cdot), u(\cdot)) = ((x_0(\cdot), x(\cdot)), u(\cdot))$ допустима в (35) и принадлежит $\mathcal{O}_1$, то пара $(x(\cdot), u(\cdot))$ допустима в задаче (**??**) и принадлежит $\mathcal{O}$. Следовательно, $x_0(t_1) = \int_{t_0}^{t_1} f(t, x(t), u(t))\,dt \geq \int_{t_0}^{t_1} f(t, \widehat{x}(t), \widehat{u}(t))\,dt = \widehat{x}_0(t_1)$ для всех $(\widehat{\overline{x}}(\cdot), \widehat{u}(\cdot)) \in \mathcal{O}_1$, т. е. $(\widehat{\overline{x}}(\cdot), \widehat{u}(\cdot))$ — слабый локальный минимум в задаче (35).

Нам понадобится следующая лемма, которую мы приводим без доказательства, поскольку она является частным случаем классического результата о дифференцируемой зависимости решения дифференциального уравнения от параметров.

Пусть $k \in \mathbb{N}$ и $\overline{u} = \overline{u}(\cdot) = (u_1(\cdot), \ldots, u_k(\cdot)) \in (C([t_0,t_1], \mathbb{R}^r))^k$. Для любого $\overline{\alpha} = (\alpha, \ldots, \alpha_k)^T \in \mathbb{R}^k$ положим $u(\cdot, \overline{\alpha}; \overline{u}) = \widehat{u}(\cdot) + \sum_{i=1}^k \alpha_i u_i(\cdot)$.

Лемма (о дифференцируемой зависимости от параметров)**.** *Пусть* $(\widehat{\overline{x}}(\cdot), \widehat{u}(\cdot))$ *— допустимая пара в задаче* (35)*, Тогда найдется такая окрестность нуля V в $\mathbb{R}^k$, что для любого $\overline{\alpha} = (\alpha_1, \ldots, \alpha_k)^T \in V$ существует единственное решение $\overline{x}(\cdot, \overline{\alpha}; \overline{u})$ задачи Коши:*

$$\dot{\overline{x}} = \overline{\varphi}(t, \overline{x}, u(t, \overline{\alpha}; \overline{u})), \quad x_0(t_0) = 0, \quad x(t_0) = x_0, \qquad (36)$$

определенное на всем отрезке $[t_0, t_1]$. Кроме того,

1) $\overline{x}(\cdot, \overline{\alpha}; \overline{u}) \to \widehat{\overline{x}}(\cdot)$ *при* $\overline{\alpha} \to 0$ *в* $C^1([t_0,t_1], \mathbb{R}^n)$*,*

2) *для каждого $t \in [t_0,t_1]$ отображение $\overline{\alpha} \mapsto \overline{x}(t, \overline{\alpha}; \overline{u})$ непрерывно дифференцируемо на V и его частные производные в нуле $\overline{x}_{\alpha_i}(t) = \overline{x}_{\alpha_i}(t, 0; \overline{u})$, $i = 1, \ldots, k$, на отрезке $[t_0, t_1]$ удовлетворяет дифференциальному уравнению*

$$\dot{\overline{x}}_{\alpha_i}(t) = \widehat{\overline{\varphi}}_x(t)\overline{x}_{\alpha_i}(t) + \widehat{\overline{\varphi}}_u(t) u_i(t), \quad \overline{x}_{\alpha_i}(t_0) = 0, \quad 1 \leq i \leq k. \qquad (37)$$

Доказательство необходимых условий минимума в задаче (P_8'). Пусть $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — слабый локальный минимум в задаче (P_8'), $k \in \mathbb{N}$, $\overline{u} = (u_1(\cdot), \ldots, u_k(\cdot)) \in (C([t_0,t_1], \mathbb{R}^r))^k$, $\overline{\alpha} = (\alpha, \ldots, \alpha_k)^T \in \mathbb{R}^k$ и $u(\cdot, \overline{\alpha}; \overline{u}) = \widehat{u}(\cdot) + \sum_{i=1}^k \alpha_i u_i(\cdot)$. Согласно лемме о дифференцируемой зависимости найдется такая окрестность нуля V в $\mathbb{R}^k$, что для любого $\overline{\alpha} = (\alpha_1, \ldots, \alpha_k)^T \in V$ существует единственное решение $\overline{x}(\cdot, \overline{\alpha}; \overline{u})$ задачи Коши (36).

Пусть $\rho > 0$ такое, что $B_{\mathbb{R}^k}(0, \rho) \subset V$. Рассмотрим задачу

$$x_0(t_1, \overline{\alpha}; \overline{u}) \to \min, \quad g(x(t_1, \overline{\alpha}; \overline{u})) = 0, \quad \overline{\alpha} \in B_{\mathbb{R}^k}(0, \rho), \qquad (38)$$

где для краткости обозначили $g = (g_1, \ldots, g_m)$.

Точка $\overline{\alpha} = 0$ доставляет локальный минимум в этой задаче. Действительно, пусть $\mathcal{O}_1$ — определенная выше окрестность точки

$\widehat{\overline{x}}(\cdot) = ((\widehat{x}_0(\cdot), \widehat{x}(\cdot))$, такая, что $x_0(t_1) \geq \widehat{x}_0(t_1)$ для любой допустимой в задаче (35) пары $(\overline{x}(\cdot), u(\cdot)) = ((x_0(\cdot), x(\cdot)), u(\cdot)) \in \mathcal{O}_1$.

По лемме о дифференцируемости зависимости от параметров $\overline{x}(\cdot, \overline{\alpha}; \overline{u}) \to \widehat{\overline{x}}(\cdot)$ при $\overline{\alpha} \to 0$ в $C^1([t_0, t_1], \mathbb{R}^{n+1})$ и, очевидно, $u(\cdot, \overline{\alpha}; \overline{u}) \to \overline{u}(\cdot)$ в $C([t_0, t_1], \mathbb{R}^r)$ при $\overline{\alpha} \to 0$. Поэтому найдется такая окрестность $U_{\mathbb{R}^k}(0, \rho_1) \subset B_{\mathbb{R}^k}(0, \rho)$, что $(\overline{x}(\cdot, \overline{\alpha}; \overline{u}), u(\cdot, \overline{\alpha}; \overline{u})) \in \mathcal{O}_1$ для всех $\overline{\alpha} \in U_{\mathbb{R}^k}(0, \rho_1)$.

Пусть вектор $\overline{\alpha}$ допустим в задаче (38) и принадлежит $U_{\mathbb{R}^k}(0, \rho_1)$. Тогда пара $(\overline{x}(\cdot, \overline{\alpha}; \overline{u}), u(\cdot, \overline{\alpha}; \overline{u}))$ допустима в задаче (35) и так как она принадлежит $\mathcal{O}_1$, то $x_0(t_1, \overline{\alpha}; \overline{u}) \geq \widehat{x}_0(t_1) = x_0(t_1, 0; \overline{u})$, т. е. $\overline{\alpha} = 0$ — локальный минимум в задаче (38).

В силу предложения 5 из §2.1 точка $\overline{\alpha} = 0$ является экстремальной для отображения $F\colon B_{\mathbb{R}^k}(0, \rho) \to \mathbb{R}^{m+1}$, где

$$F(\overline{\alpha}) = F(\overline{\alpha}; \overline{u}) = (x_0(t_1, \overline{\alpha}; \overline{u}),\ g(x(t_1, \overline{\alpha}; \overline{u}))^T.$$

Согласно основной теореме найдется ненулевой вектор $\lambda(\overline{u}) \in (\mathbb{R}^{m+1})^*$ такой, что

$$\langle \lambda(\overline{u}), F'(0; \overline{u})\overline{\alpha} \rangle \geq 0, \quad \forall\, \overline{\alpha} \in B_{\mathbb{R}^k}(0, \rho).$$

Как видно из (37), частная производная отображения $\overline{\alpha} \mapsto \overline{x}(t_1, \overline{\alpha}; \overline{u})$ по α_i, $1 \leq i \leq k$, в нуле зависит только от функции $u_i(\cdot)$. Поэтому и частная производная отображения F по α_i в нуле также зависит только от $u_i(\cdot)$. Обозначим ее $F_{\alpha_i}(0; u_i(\cdot))$. Тогда последнее неравенство, учитывая, что оно, очевидно, верно для любых $\overline{\alpha} \in \mathbb{R}^k$, запишется в виде следующего равенства

$$\langle \lambda(\overline{u}), \sum_{i=1}^{k} F_{\alpha_i}(0; u_i(\cdot))\alpha_i \rangle = 0, \quad \forall\, \overline{\alpha} = (\alpha_1, \ldots, \alpha_k)^T \in \mathbb{R}^k. \tag{39}$$

Можно считать, что $|\lambda(\overline{u})| = 1$. Обозначим множество всех таких $\lambda(\overline{u})$, удовлетворяющих (39), через $\Lambda_k(\overline{u})$. Ясно, что $\Lambda_k(\overline{u})$ — замкнутое подмножество компакта — единичной сферы в $(\mathbb{R}^{m+1})^*$.

Таким образом, каждому конечному набору функций из $C([t_0, t_1], \mathbb{R}^r)$ можно сопоставить замкнутое подмножество указанного компакта. Покажем, что семейство $\mathcal{A}$ всех таких подмножеств образует центрированную систему, т. е. любое конечное его подсемейство имеет непустое пересечение (см. п. 1.1.2).

Пусть $\Lambda_{k_j}(\overline{u}_j)$, $1 \leq j \leq s$, ($\overline{u}_j$ — набор из k_j функций) — произвольное конечное семейство семейства $\mathcal{A}$. Проверим, что $\cap_{j=1}^{s} \Lambda_{k_j}(\overline{u}_j) \neq \emptyset$. Рассмотрим набор: $\widetilde{u} = \cup_{j=1}^{s} \overline{u}_j$, состоящий из объединения множеств функций, входящих в наборы $\overline{u}_j$, $j = 1, \ldots, s$, и пусть $\lambda(\widetilde{u}) \in \Lambda_l(\widetilde{u})$, где l — число функций в наборе $\widetilde{u}$. Пусть

$1 \le j \le s$. В соотношении (39) (выписанным для набора $\widetilde{u}$) положим $\alpha_i = 0$ для тех индексов i, для которых $u_i(\cdot)$ не принадлежит набору $\overline{u}_j$. Тогда получим, что $\lambda(\widetilde{u}) \in \Lambda_{k_j}(\overline{u}_j)$, т. е. $\lambda(\widetilde{u}) \in \cap_{j=1}^{s}\Lambda_{k_j}(\overline{u}_j)$.

Таким образом, совокупность множеств $\mathcal{A}$ образует центрированную систему и значит, согласно лемме о центрированной системе из §1.2 существует λ, для которого справедливо соотношение (39) для любого $k \in \mathbb{N}$ и любого набора $\overline{u}$. В частности, оно справедливо для наборов, состоящих из одной функции, т. е. $k = 1$ и $\overline{u} = u_1(\cdot)$. В этом случае $\overline{\alpha} = \alpha_1$. Будем писать $u(\cdot)$ вместо $u_1(\cdot)$ и α вместо α_1 Полагая $\alpha = 1$, соотношение (39) в этой ситуации перейдет в равенство

$$\langle \lambda, F_\alpha(0; u(\cdot))\rangle = 0, \quad \forall\, u(\cdot) \in C([t_0, t_1], \mathbb{R}^r).$$

Отображение $\alpha \mapsto F(\alpha; u(\cdot))$ есть суперпозиция отображения $\alpha \mapsto \overline{x}(t_1, \alpha; u(\cdot))$ и отображения $\overline{g}$, сопоставляющее вектору $(x_0, x_1, \dots, x_n)^T$ вектор $(x_0, g(x))^T$. Учитывая это, последнее равенство запишется так (как и в лемме, для краткости, пишем $\overline{x}_\alpha(t)$ вместо $\overline{x}_\alpha(t, 0; u(\cdot))$)

$$\langle \lambda, \overline{g}'(\widehat{\overline{x}}(t_1))\overline{x}_\alpha(t_1)\rangle = 0, \quad \forall\, u(\cdot) \in C([t_0, t_1], \mathbb{R}^r). \tag{40}$$

Пусть $\overline{p} = (p_0, p) = (p_0, p_1, \dots, p_n)$ — решение линейного дифференциального уравнения

$$-\dot{\overline{p}}(t) = \overline{p}(t)\,\widehat{\overline{\varphi}}_x(t), \quad \overline{p}(t_1) = -\lambda\overline{g}'(\widehat{\overline{x}}(t_1)). \tag{41}$$

Умножая при каждом $t \in [t_0, t_1]$ обе части этого уравнения на $\overline{x}_\alpha(t)$, а затем используя дифференциальное уравнение в (37), получим для всех $t \in [t_0, t_1]$ тождество $\langle \dot{\overline{p}}(t), \overline{x}_\alpha(t)\rangle + \langle \overline{p}(t), \dot{\overline{x}}_\alpha(t)\rangle = \langle \overline{p}(t)\widehat{\overline{\varphi}}_u(t), u(t)\rangle$, интегрируя которое, учитывая начальные условия в (37) и (41), будем иметь в силу (40)

$$0 = \langle \lambda, \overline{g}'(\widehat{\overline{x}}(t_1))\overline{x}_\alpha(t_1)\rangle = \langle \lambda, \overline{g}'(\widehat{\overline{x}}(t_1))\overline{x}_\alpha(t_1)\rangle + \int_{t_0}^{t_1} (\langle \dot{\overline{p}}(t), \overline{x}_\alpha(t)\rangle$$
$$+ \langle \overline{p}(t), \dot{\overline{x}}_\alpha(t)\rangle)\, dt - \int_{t_0}^{t_1} \langle \overline{p}(t)\widehat{\overline{\varphi}}_u(t), u(t)\rangle\, dt = \langle \lambda, \overline{g}'(\widehat{\overline{x}}(t_1))\overline{x}_\alpha(t_1)\rangle$$
$$+ \langle \overline{p}(t), \overline{x}_\alpha(t)\rangle|_{t_0}^{t_1} - \int_{t_0}^{t_1} \langle \overline{p}(t)\widehat{\overline{\varphi}}_u(t), u(t)\rangle\, dt = -\int_{t_0}^{t_1} \langle \overline{p}(t)\widehat{\overline{\varphi}}_u(t), u(t)\rangle\, dt.$$

Это равенство справедливо для всех $u(\cdot) \in C([t_0, t_1], \mathbb{R}^r)$, поэтому полагая в интеграле справа $u(\cdot) = \overline{p}(\cdot)\widehat{\overline{\varphi}}_u(\cdot)$, получаем, что

$$\overline{p}(t)\widehat{\overline{\varphi}}_u(t) = 0, \quad \forall\, t \in [t_0, t_1]. \tag{42}$$

Пусть $\lambda = (\lambda_0, \overline{\lambda}) \in \mathbb{R} \times (\mathbb{R}^m)^*$. Тогда (41) равносильно тому, что

$$-\dot{p}_0(t) = 0, \quad -\dot{p}(t) = p(t)\widehat{\varphi}_x(t) + p_0\widehat{f}_{0x}(t), \quad p_0(t_1) = -\lambda_0,$$
$$p(t_1) = -\overline{\lambda}g'(\widehat{x}(t_1)). \quad (43)$$

Ясно, что $p(\cdot) \in C^1([t_0,t_1],(\mathbb{R}^n)^*)$.

Равенство (42) равносильно соотношению

$$p(t)\widehat{\varphi}_u(t) - \lambda_0\widehat{f}_{0u}(t) = 0, \quad \forall\, t \in [t_0,t_1].$$

Отсюда, из (43) и вида лагранжиана L следуют все утверждения теоремы. □

§ 4.4. Задача оптимального управления.

4.4.1. Постановка задачи.

Пусть $[t_0,t_1]$ — конечный отрезок, G — открытое подмножество $\mathbb{R}\times\mathbb{R}^n$, U — непустое подмножество $\mathbb{R}^r$, W — открытое подмножество $\mathbb{R}^n\times\mathbb{R}^n$, заданы функции $f_i\colon G\times U \to \mathbb{R}$, $i = 0,1,\dots,m$, и отображение $\varphi\colon G\times U \to \mathbb{R}^n$ 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 $l_i\colon W \to \mathbb{R}$, $i = 0,1,\dots,m$ переменных ξ_0 и ξ_1.

Обозначим через $PC^1([t_0,t_1],\mathbb{R}^n)$ совокупность всех кусочно-непрерывно дифференцируемых, а через $PC([t_0,t_1],\mathbb{R}^r)$ совокупность кусочно-непрерывных функций на $[t_0,t_1]$ со значениями соответственно в $\mathbb{R}^n$ и $\mathbb{R}^r$.

Пусть $\mathcal{B}_i(x(\cdot),u(\cdot))$, $i = 0,1,\dots,m$, обозначают тоже, что и в п. 4.3.1. В пространстве $Z = PC^1([t_0,t_1],\mathbb{R}^n)\times PC([t_0,t_1],\mathbb{R}^r)$ относительно переменной $(x(\cdot),u(\cdot)) \in Z$ рассмотрим задачу

$$\mathcal{B}_0(x(\cdot),u(\cdot)) \to \min, \quad \dot{x}(t) = \varphi(t,x(t),u(t)), \quad u(t)\in U,$$
$$\mathcal{B}_i(x(\cdot),u(\cdot)) \le 0, \quad 1 \le i \le m',$$
$$\mathcal{B}_i(x(\cdot),u(\cdot)) = 0, \quad m'+1 \le i \le m, \quad (P_9)$$

причем равенство $\dot{x}(t) = \varphi(t,x(t),u(t))$ и включение $u(t)\in U$ должны выполняться для всех $t\in[t_0,t_1]$, где функция $u(\cdot)$ непрерывна.

Задача (P_9) называется *задачей оптимального управления*. Переменную $x(\cdot)$ часто называют фазовой переменной, а $u(\cdot)$ — управлением.

Пара $(x(\cdot),u(\cdot)) \in Z$ называется *допустимым процессом* в задаче (P_9), если $\Gamma(x(\cdot)) = \{\,(t,x(t)) \mid t\in[t_0,t_1]\,\} \subset G$, $(x(t_0),x(t_1)) \in W$ и $(x(\cdot),u(\cdot))$ удовлетворяет ограничениям задачи.

Допустимый процесс $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ называется *оптимальным процессом* или *сильным минимумом* в задаче (P_9), если существует такая окрестность $\mathcal{O}$ функции $\widehat{x}(\cdot)$ в $C([t_0, t_1], \mathbb{R}^n)$, что для любого допустимого процесса $(x(\cdot), u(\cdot))$, для которого $x(\cdot) \in \mathcal{O}$, выполнено неравенство $\mathcal{B}_0(x(\cdot), u(\cdot)) \geq \mathcal{B}_0(\widehat{x}(\cdot), \widehat{u}(\cdot))$.

4.4.2. Принцип максимума Понтрягина - необходимые условия минимума первого порядка.

Положим $H(t, x, u, \lambda, p) = \langle p, \varphi(t, x, u)\rangle - \sum_{i=0}^{m} \lambda_i f_i(t, x, u)$ (*функцию Понтрягина* задачи (P_9)), $l(\xi_0, \xi_1) = \sum_{i=0}^{m} \lambda_i l_i(\xi_0, \xi_1)$, где $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и $p \in (\mathbb{R}^n)^*$.

Если фиксирована пара $(\widehat{x}(\cdot), \widehat{u}(\cdot))$, то для функций и отображений на этой паре пользуемся теми же сокращенными обозначениями, что и в задаче Лагранжа. Если еще фиксирована функция $p = p(\cdot)$ на $[t_0, t_1]$ со значениями в $(\mathbb{R}^n)^*$, то пишем $\widehat{H}(t, p) = \langle p(t), \widehat{\varphi}(t)\rangle - \sum_{i=0}^{m} \lambda_i \widehat{f}_i(t)$ и аналогично для производной по x этой функции.

Теорема (Принцип максимума Понтрягина)**.** *Пусть $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — оптимальный процесс в задаче (P_9). Если функции $f_i\colon G \times \mathrm{cl}\, U \to \mathbb{R}$, $0 \leq i \leq m$, и отображение $\varphi\colon G \times \mathrm{cl}\, U \to \mathbb{R}^n$ непрерывны вместе со своими частными производными по x на $G \times \mathrm{cl}\, U$, а функции l_i, $0 \leq i \leq m$, непрерывно дифференцируемы на W, то найдутся вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in (\mathbb{R}^{m+1})^*$ и функция $p(\cdot) \in PC^1([t_0, t_1], (\mathbb{R}^n)^*)$, не равные одновременно нулю, такие, что выполняются*

(a) *условие стационарности по x*

$$\dot{p}(t) = -\widehat{H}_x(t, p) \quad \Leftrightarrow \quad -\dot{p}(t) = p(t)\widehat{\varphi}_x(t) - \sum_{i=0}^{m} \lambda_i \widehat{f}_{ix}(t),$$

(b) *условия трансверсальности*

$$p(t_i) = (-1)^i \widehat{l}_{\xi_i}, \quad i = 0, 1,$$

(c) *условие максимума по u для всех $t \in [t_0, t_1]$, где функция $\widehat{u}(\cdot)$ непрерывна*

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

(d) *условия дополняющей нежесткости*

$$\lambda_i \mathcal{B}_i(\widehat{x}(\cdot), \widehat{u}(\cdot)) = 0, \quad i = 1, \ldots, m',$$

(e) *условия неотрицательности*

$$\lambda_i \geq 0, \quad i = 0, 1, \ldots m'.$$

Из-за специфического условия (c) необходимые условия в задаче оптимального управления обычно называют “Принципом максимума Понтрягина”.

Как это было сделано в случае задачи Лагранжа, доказательство проведем для частного случая постановки (P_9), а именно, рассмотрим такую задачу

$$\int_{t_0}^{t_1} f_0(t,x(t),u(t))\,dt \to \min, \quad \dot{x}(t) = \varphi(t,x(t),u(t)), \quad u(t) \in U,$$
$$x(t_0) = x_0, \quad g_i(x(t_1)) = 0, \quad i = 1,\dots,m, \quad (P_9')$$

где g_i, $i = 1,\dots,m$, — непрерывно дифференцируемые функции на некотором открытом множестве в $\mathbb{R}^n$.

В данном случае $H(t,x,u,\lambda_0,p) = \langle p, \varphi(t,x,u)\rangle - \lambda_0 f_0(t,x,u)$, $l(x(t_0),x(t_1)) = \langle \mu, x(t_0)\rangle + \sum_{i=1}^m \lambda_i g_i(x(t_1))$, где $\mu \in (\mathbb{R}^n)^*$, $\lambda = (\lambda_1,\dots,\lambda_m) \in (\mathbb{R}^m)^*$ и $p \in (\mathbb{R}^n)^*$.

Тогда утверждения теоремы для задачи (P_9'), как легко проверить, будут таковы: утверждения (a) и (c) останутся прежними, но с $m = 0$, условия дополняющей нежесткости отсутствуют, $\lambda_0 \geq 0$, а условия трансверсальности имеют вид

$$p(t_1) = -\sum_{i=1}^m \lambda_i g_i'(\widehat{x}(t_1)).$$

Основания для того, чтобы не писать еще условие $p(t_0) = \mu$ те же, что и в соответствующей задаче Лагранжа.

Задачу (P_9'), аналогично тому как это было сделано в задаче Лагранжа, приведем к задаче без интегрального функционала. Для этого, как и раньше, обозначим $\varphi_0 = f_0$, $\overline{\varphi} = (\varphi_0, \varphi)^T = (\varphi_0, \varphi_1, \dots, \varphi_n)^T$, $\overline{x} = (x_0, x)^T = (x_0, x_1, \dots, x_n)^T$ и рассмотрим задачу

$$x_0(t_1) \to \min, \quad \dot{\overline{x}} = \overline{\varphi}(t,\overline{x},u), \quad u(t) \in U,$$
$$x_0(t_0) = 0, \quad x(t_0) = x_0, \quad g_i(x(t_1)) = 0, \quad i = 1,\dots,m. \quad (44)$$

Заметим, что если $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ — оптимальный процесс в задаче (P_9'), то $(\widehat{\overline{x}}(\cdot), \widehat{u}(\cdot)) = ((\widehat{x}_0(\cdot), \widehat{x}(\cdot)), \widehat{u}(\cdot))$, где $\widehat{x}_0(t) = \int_{t_0}^t f(\tau, \widehat{x}(\tau), \widehat{u}(\tau))\,d\tau$, — оптимальный процесс в задаче (44). Проверка этого факта также аналогично тому, как это было сделано в задаче Лагранжа.

Перед непосредственным доказательством теоремы приведем некоторые определения и сформулируем одно утверждение.

Пусть $\widehat{u}(\cdot)$ — кусочно непрерывная функция на отрезке $[t_0, t_1]$, τ — точка непрерывности $\widehat{u}(\cdot)$ и $v \in U$. Для всех $\alpha \geq 0$ таких, что

$\widehat{u}(\cdot)$ непрерывна на отрезке $[\tau-\alpha,\tau]$ определим на $[t_0,t_1]$ кусочно непрерывную функцию

$$u(t,\alpha;\tau,v)=\begin{cases}\widehat{u}(t), & \text{если} \quad t\notin[\tau-\alpha,\tau);\\ v, & \text{если} \quad t\in[\tau-\alpha,\tau),\end{cases}$$

которую называют *игольчатой вариацией* функции $\widehat{u}(\cdot)$, а пару (τ,v) — *иголкой*. При $\alpha=0$ игольчатая вариация совпадает с $\widehat{u}(\cdot)$.

Понятие игольчатой вариации естественным образом распространяется на конечный набор иголок. Пусть $k\in\mathbb{N}$ и $\mathcal{N}_k$ — совокупность пар (иголок) $(\tau_i,v_i)\in[t_0,t_1]\times U$, $1\le i\le k$, где $t_0<\tau_1<\ldots<\tau_k<t_1$ и точки τ_i суть точки непрерывности $\widehat{u}(\cdot)$. Для каждого $\overline{\alpha}=(\alpha_1,\ldots,\alpha_k)^T\in\mathbb{R}^k_+$ такого, что отрезки $[\tau_i-\alpha_i,\tau_i]$, $1\le i\le k$, не пересекаются и на них $\widehat{u}(\cdot)$ непрерывна, определим на $[t_0,t_1]$ кусочно непрерывную функцию

$$u(t,\overline{\alpha};\mathcal{N}_k)=\begin{cases}\widehat{u}(t), & \text{если} \quad t\notin\cup_{i=1}^k[\tau_i-\alpha_i,\tau_i);\\ v_i, & \text{если} \quad t\in[\tau_i-\alpha_i,\tau_i),\ 1\le i\le k,\end{cases}$$

которая также называется *игольчатой вариацией* $\widehat{u}(\cdot)$, а набор $\mathcal{N}_k$ называется *пакетом иголок*.

Лемма (о пакете иголок)**.** *Пусть* $(\widehat{\overline{x}}(\cdot),\widehat{u}(\cdot))$ *— допустимая пара в задаче* (44), $k\in\mathbb{N}$ *и* $\mathcal{N}_k=\{(\tau_i,v_i),\ 1\le i\le k\}$ *— пакет иголок. Тогда найдется такая окрестность нуля* V *в* $\mathbb{R}^k$, *что для любого* $\overline{\alpha}\in V\cap\mathbb{R}^k_+$ *существует единственное решение* $\overline{x}(\cdot,\overline{\alpha};\mathcal{N}_k)$ *задачи Коши*

$$\dot{\overline{x}}=\overline{\varphi}(t,\overline{x},u(t,\overline{\alpha};\mathcal{N}_k)),\quad x_0(t_0)=0,\quad x(t_0)=x_0, \tag{45}$$

определенное на всем отрезке $[t_0,t_1]$. *При этом,*

1) $\overline{x}(\cdot,\overline{\alpha};\mathcal{N}_k)\to\widehat{x}(\cdot)$ *при* $\overline{\alpha}\to 0$ *равномерно на* $[t_0,t_1]$,

2) *отображение* $\overline{\alpha}\mapsto\overline{x}(t_1,\overline{\alpha};\mathcal{N}_k)$ *строго дифференцируемо в нуле* (*относительно* $\mathbb{R}^k_+$), *а частная производная по* α_i, $1\le i\le k$, *в нуле отображения* $\overline{\alpha}\mapsto\overline{x}(t,\overline{\alpha};\mathcal{N}_k)$, *которую обозначим* $\overline{x}_{\alpha_i}(t)=\overline{x}_{\alpha_i}(t,0;\mathcal{N}_k)$, *на* $[\tau_i,t_1]$ *удовлетворяет уравнению*

$$\dot{\overline{x}}_{\alpha_i}(t)=\widehat{\overline{\varphi}}_x(t)\overline{x}_{\alpha_i}(t),\ \overline{x}_{\alpha_i}(\tau_i)=\widehat{\overline{\varphi}}(\tau_i,\widehat{\overline{x}}(\tau_i),v_i)-\widehat{\overline{\varphi}}(\tau_i,\widehat{\overline{x}}(\tau_i),\widehat{u}(\tau_i)). \tag{46}$$

Доказательство принципа максимума для задачи (P_9').
Пусть $(\widehat{x}(\cdot),\widehat{u}(\cdot))$ — оптимальный процесс в задаче (P_9'), $k\in\mathbb{N}$, $\mathcal{N}_k=\{(\tau_i,v_i),\ 1\le i\le k\}$ — пакет иголок и $u(t,\overline{\alpha};\mathcal{N}_k)$ — соответствующая игольчатая вариация управления $\widehat{u}(\cdot)$. Согласно лемме о пакете иголок найдется такая окрестность нуля V в $\mathbb{R}^k$, что для любого $\overline{\alpha}\in V\cap\mathbb{R}^k_+$ существует единственное решение $\overline{x}(\cdot,\overline{\alpha};\mathcal{N}_k)$ задачи Коши (45).

Пусть $\rho > 0$ такое, что $B_{\mathbb{R}^k}(0,\rho) \subset V$. Рассмотрим задачу относительно переменной $\overline{\alpha}$:

$$x_0(t_1,\overline{\alpha};\mathcal{N}_k) \to \min,\ \ g(x(t_1,\overline{\alpha};\mathcal{N}_k)) = 0,\ \ \overline{\alpha} \in B_{\mathbb{R}^k}(0,\rho) \cap \mathbb{R}^k_+, \quad (47)$$

где $g = (g_1,\ldots,g_m)$.

Точка $\overline{\alpha} = 0$ доставляет локальный минимум в этой задаче. Проверка этого факта, фактически, такая же, как и для задачи Лагранжа.

В силу предложения 5 из §2.1 точка $\overline{\alpha} = 0$ является экстремальной для отображения $F\colon B_{\mathbb{R}^k}(0,\rho) \cap \mathbb{R}^k_+ \to \mathbb{R}^{m+1}$, где

$$F(\overline{\alpha}) = F(\overline{\alpha};\mathcal{N}_k) = (x_0(t_1,\overline{\alpha};\mathcal{N}_k),\ g(x(t_1,\overline{\alpha};\mathcal{N}_k)))^T.$$

Тогда согласно основной теореме найдется ненулевой вектор $\lambda(\mathcal{N}_k) \in (\mathbb{R}^{m+1})^*$ такой, что

$$\langle \lambda(\mathcal{N}_k), F'(0;\mathcal{N}_k)\overline{\alpha}\rangle \geq 0, \quad \forall\ \overline{\alpha} \in B_{\mathbb{R}^k}(0,\rho) \cap \mathbb{R}^k_+,$$

где $F'(0)$ — производная в нуле отображения F относительно конуса $\mathbb{R}^k_+$.

Как видно из (46), частная производная по α_i, $1 \leq i \leq k$, в нуле отображения $\overline{\alpha} \mapsto \overline{x}(t_1,\overline{\alpha};\mathcal{N}_k)$ зависит только от иголки (τ_i,v_i), поэтому и частная производная по α_i, $1 \leq i \leq k$, в нуле отображения F зависит только от (τ_i,v_i). Обозначим ее $F_{\alpha_i}(0;\tau_i,v_i)$. Тогда последнее неравенство, учитывая, что оно, очевидно, верно для любых $\overline{\alpha} \in \mathbb{R}^k_+$, запишется следующим образом

$$\langle \lambda(\mathcal{N}_k), \sum_{i=1}^k F_{\alpha_i}(0;\tau_i,v_i)\alpha_i\rangle \geq 0, \quad \forall\ \overline{\alpha} = (\alpha_1,\ldots,\alpha_k)^T \in \mathbb{R}^k_+. \quad (48)$$

Можно считать, что $|\lambda(\mathcal{N}_k)| = 1$. Обозначим множество всех таких $\lambda(\mathcal{N}_k)$, удовлетворяющих (48), через $\Lambda(\mathcal{N}_k)$. Ясно, что $\Lambda(\mathcal{N}_k)$ — замкнутое подмножество компакта — единичной сферы в $(\mathbb{R}^{m+1})^*$. Таким образом, каждому $k \in \mathbb{N}$ и каждому набору $\mathcal{N}_k$ можно сопоставить замкнутое подмножество указанного компакта. Покажем, что семейство $\mathcal{A}$ всех таких подмножеств образует центрированную систему.

Пусть $\Lambda(\mathcal{N}_{k_j})$, $1 \leq j \leq s$, — произвольное конечное семейство семейства $\mathcal{A}$. Проверим, что $\cap_{j=1}^s \Lambda(\mathcal{N}_{k_j}) \neq \emptyset$. Положим $\mathcal{N}_l = \cup_{j=1}^s \mathcal{N}_{k_j}$ (т. е. $\mathcal{N}_l$ — объединение иголок из пактов $\mathcal{N}_{k_j}$, $1 \leq j \leq s$, и l — число элементов в этом объединении). Если некоторые точки τ_i совпали, то рассмотрим точки $t_0 < \tau_{1p} < \ldots < \tau_{lp} < t_1$ такие, что $\tau_{ip} \to \tau_i$ при $p \to \infty$. Для пакета $\mathcal{N}_{l(p)} = \{(\tau_{ip},v_i),\ 1 \leq i \leq l\}$ соотношение (48) справедливо с некоторым $\lambda(\mathcal{N}_{l(p)})$, $|\lambda(\mathcal{N}_{l(p)})| = 1$. Но $x_{\alpha_i}(t_1;\tau_{ip},v_i) \to x_{\alpha_i}(t_1;\tau_i,v_i)$ при $p \to \infty$ в силу непрерывной зависимости решения дифференциального уравнения (см. (46)) от начальных данных. Можно считать, что $\lambda(\mathcal{N}_{l(p)})$ также сходится к некоторому $\lambda(\mathcal{N}_l)$ при $p \to \infty$ и таким образом, соотношение (48) справедливо для набора $\mathcal{N}_l$ с $\lambda(\mathcal{N}_l)$.

Пусть $1 \le j \le s$. Положим в этом соотношении $\alpha_i = 0$ для тех индексов i, для которых (τ_i, v_i) не принадлежит набору $\mathcal{N}_{k_j}$. Тогда получим, что $\lambda(\mathcal{N}_l) \in \Lambda(\mathcal{N}_{k_j})$ и тем самым $\lambda(\mathcal{N}_l) \in \cap_{j=1}^{s} \Lambda(\mathcal{N}_{k_j})$.

Итак, совокупность множеств $\mathcal{A}$ образует центрированную систему и значит, согласно лемме о центрированной системе из §1.2 существует λ, для которого справедливо соотношение (48) при любом $k \in \mathbb{N}$ и любом наборе $\mathcal{N}_k$. В частности, для любого набора $\mathcal{N}_1 = \{(\tau, v)\}$, где $(\tau, v) \in (t_0, t_1) \times U$ и τ — точка непрерывности $\widehat{u}(\cdot)$ следует из (48), полагая $\alpha = 1$, неравенство

$$\langle \lambda, F_\alpha(0; \tau, v) \rangle \ge 0.$$

Как и в задаче Лагранжа, отображение $\alpha \mapsto F(\alpha; \tau, v)$ есть суперпозиция отображения $\alpha \mapsto \overline{x}(t_1, \alpha; \tau, v)$ и отображения $\overline{g}$, сопоставляющее вектору $(x_0, x_1, \ldots, x_n)^T$ вектор $(x_0, g(x))^T$. Учитывая это, последнее неравенство запишется так (для краткости пишем $\overline{x}_\alpha(t_1)$ вместо $\overline{x}_\alpha(t_1, 0; \tau, v)$)

$$\langle \lambda, \overline{g}'(\widehat{\overline{x}}(t_1))\overline{x}_\alpha(t_1) \rangle \ge 0. \tag{49}$$

Пусть $\overline{p} = (p_0, p) = (p_0, p_1, \ldots, p_n)$ — решение линейного дифференциального уравнения

$$-\dot{\overline{p}}(t) = \overline{p}(t)\,\widehat{\overline{\varphi}}_x(t), \quad \overline{p}(t_1) = -\lambda \overline{g}'(\widehat{\overline{x}}(t_1)). \tag{50}$$

Умножая при каждом $t \in [\tau, t_1]$ обе части этого уравнения на $\overline{x}_\alpha(t)$, а затем используя дифференциальное уравнение в (46) (где вместо τ_i и v_i надо подставить соответственно τ и v), получим тождество $\langle \dot{\overline{p}}(t), \overline{x}_\alpha(t) \rangle + \langle \overline{p}(t), \dot{\overline{x}}_\alpha(t) \rangle = 0$ для всех $t \in [\tau, t_1]$ Учитывая это, начальные условия в (46) и в (50), будем иметь из (49)

$$0 \le \langle \lambda, \overline{g}'(\widehat{\overline{x}}(t_1))\overline{x}_\alpha(t_1) \rangle = \langle \lambda, \overline{g}'(\widehat{\overline{x}}(t_1))\overline{x}_\alpha(t_1) \rangle + \int_\tau^{t_1} (\langle \dot{\overline{p}}(t), \overline{x}_\alpha(t) \rangle$$
$$+ \langle \overline{p}(t), \dot{\overline{x}}_\alpha(t) \rangle)\, dt = \langle \lambda, \overline{g}'(\widehat{\overline{x}}(t_1))\overline{x}_\alpha(t_1) \rangle + \langle \overline{p}(t), \overline{x}_\alpha(t) \rangle |_\tau^{t_1}$$
$$= -\langle \overline{p}(\tau), \widehat{\overline{\varphi}}(\tau, \widehat{x}(\tau), v) - \widehat{\overline{\varphi}}(\tau, \widehat{x}(\tau), \widehat{u}(\tau) \rangle. \tag{51}$$

Если $\lambda = (\lambda_0, \overline{\lambda}) \in \mathbb{R} \times (\mathbb{R}^m)^*$, то дифференциальное уравнение в (50) равносильно тому, что

$$-\dot{p}_0(t) = 0, \quad -\dot{p}(t) = p(t)\widehat{\varphi}_x(t) + p_0 \widehat{f}_{0x}(t), \quad p_0(t_1) = -\lambda_0,$$
$$p(t_1) = -\overline{\lambda} g'(\widehat{x}(t_1)). \tag{52}$$

Ясно, что $p(\cdot) \in PC^1([t_0, t_1], (\mathbb{R}^n)^*)$.

Подставляя в (51) $(-\lambda_0, p(\cdot))$ вместо $\overline{p}(\cdot)$ и $(\widehat{f}(\cdot), \widehat{\varphi}(\cdot))$ вместо $\widehat{\widehat{\varphi}}(\cdot)$, получим, что $H(\tau, \widehat{x}(\tau), v, \lambda_0, p(\tau)) \leq H(\tau, \widehat{x}(\tau), \widehat{u}(\tau), \lambda_0, p(\tau))$ для всех точек непрерывности τ функции $\widehat{u}(\cdot)$ и $v \in U$. Отсюда и (52) следуют все утверждения теоремы. $\square$