КУРС ЛЕКЦИЙ "ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ И ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ" (4-Й КУРС)

А. В. Фурсиков

Лекция 1

1. Примеры экстремальных задач

1.1. **Задача Дидоны.** (825г до н.э.: Основание Карфагена)

Постановка задачи: Пусть Ω -ограниченная область на двумерной плоскости с гладкой границей $\partial\Omega$. При заданной длине границы $L = L(\partial\Omega)$ найти область Ω с максимальной площадью $S = S(\Omega)$

Ответ: Ω - это круг.

Полезная формула: $L^2(\partial\Omega) \geq 4\pi S(\Omega)$

1.2. **Задача Апполония.** (III-II век до н.э. из книги "Коника")

Постановка задачи: Даны эллипс в плоскости $\mathbb{R}^2$ и точка A вне него. Найти расстояние от A до эллипса.

Эквивалентная постановка задачи: Пусть $A = (x_0, y_0)$. Решить экстремальную задачу:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$
$$(x - x_0)^2 + (y - y_0)^2 \to \min$$

1.3. **Задача о брахистохроне.** (И.Бернулли 1696г. Задача вариационного исчисления)

Постановка задачи: В вертикальной плоскости даны точки A и B. Пусть A находится выше B. Найти такую кривую, соединяющую A, B, по которой точка скатывается от A до B под действием силы тяжести за кратчайшее время (трение не учитывается.)

Дадим математическую постановку задачи. Пусть A - начало координат, из которого выходят горизонтальный и вертикальный (направленный вниз) элементы базиса, задающие исходную вертикальную плоскость. Любая точка этой плоскости имеет координаты (x, y) , где x определяется горизонтальным, а y-вертикальным элементами базиса. Будем искать кривую, соединяющую точки A, B, в виде $(x, y(x))$. Покажем, что любая кривая такого вида удовлетворяет следующему принципу Галлилея: скорость $v(x)$ материальной точки, двигающейся при свободном падении вдоль этой кривой определяется по формуле

$$v(x) = \sqrt{2gy(x)},$$

где g - ускорение свободного падения.

Действительно, при свободном падении $y''(t) = g$ и значит $y(t) = a + b(t - t_0) + c(t - t_0)^2$. Если в момент времени $t_0 y = 0$, то $a = 0, b = 0$ и значит $g = 2c$ то есть $c = g/2$. Отсюда следует, что $v(t) = y'(t) = g(t - t_0) = \sqrt{2g\frac{g}{2}(t - t_0)^2}$. Отсюда следует принцип Галлилея.

Решение задачи о брахистохроне состоит в минимизации интеграла

$$T = \int_{AB} \frac{ds}{v} = \int_0^{x_1} \frac{\sqrt{1+(y'(x))^2}}{\sqrt{2gy(x)}} dx$$

при условии, что

$$y(0) = 0, \quad y(x_1) = y_1$$

1.4. **Транспортная задача.** (Л.В.Канторович-основоположник мат.экономики и лауреат нобелевской премии. 30-е годы XX века)

Постановка задачи: Имеется nскладов и m магазинов. В i-м складе есть a_i товара. Надо в j-й магазин доставить b_j товара. Стоимость транспортировки единицы товара из i-го склада в j магазин равна d_{ij}, а количество товара, которое необходимо перевезти из i-го склада в j-й магазин равно x_{ij}. Найти такое x_{ij}, чтобы стоимость транспортировки была наименьшей.

Математическая формулировка задачи:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} d_{ij} x_{ij} \to \min$$

$$\sum_{j=1}^{m} x_{ij} \leq a_j$$

$$\sum_{i=1}^{n} x_{ij} = b_j$$

Задачи такого типа называются задачами линейного программирования.

1.5. **Задача быстродействия.** Задачи такого типа возникли в 50-х годах XX-го века в связи с потребносью в космонавтике.

Математическая формулировка задачи:

$$x''(t) = u(t), \quad |u(t)| \leq 1, x(o) = x_0, \quad x'(0) = x_1, \quad x(T) = 0$$

$$T \to \min$$

Задача является простейшей задачей оптимального управления. Основоположником этого направления является Л.С.Понтрягин.

1.6. **Задача Ньютона.** Изучена И.Ньютоном в его книге "Математические задачи натуральной философии". Посвящена вопросу о сопротивлении тела, движущегося в "редкой среде"(в газе) Ньютон рассмотрел случаи шара, цилиндра, конуса.

$$F = k \int_0^R \frac{rdr}{1+(dx/dr)^2} \to \min, \quad x(0) = 0, \quad x(R) = \xi, \quad dx/dr \in \mathbb{R}_+$$

1.7. **Литература по курсу.** При чтении этих лекций предполагается использовать следующую литературу:

[1] В.М.Алексеев, В.М.Тихомиров, С.В.Фомин. "Оптимальное управление". Москва "Наука"1979г.

[2] Э.М.Галеев, М.И.Зеликин, С.В.Конягин и др. "Оптимальное управление". Москва, Издательство МЦНМО, 2008

[3] В.М.Алексеев, Э.М.Галеев, В.М.Тихомиров. "Сборник задач по оптимизации". Москва "Наука"1984г.

[4] Э.М.Галеев. "Сборник задач по вариационному исчислению и оптимальному управлению". Москва, МГУ, 2020г.

В частности, в учебнике [1] можно найти более подробное описание задач, рассмотренных выше, а также большое количество других аналогичных задач.

Лекция 2

2. Простейшая задача классического вариационного исчисления

2.1. **Постановка задачи.** В пространстве $C^1[t_0,t_1]$ непрерывно дифференцируемых функций, заданных на отрезке $[t_0,t_1]$, рассматривается следующая задача:

$$J(x)=\int_{t_0}^{t_1} L(t,x(t),\dot{x}(t))dt \to \text{extr}, \tag{1}$$

$$x(t_0)=x_0, \quad x(t_1)=x_1, \tag{2}$$

где $L(t,x,y)\in C^1([t_0,t_1]\times\mathbb{R}\times\mathbb{R})$-заданная функция, искомой является функция $x(t)\in C^1[t_0,t_1]$, а x_0,x_1 - заданные вещественные числа.

Определение 2.1. *Функция $x(t)\in C^1[t_0,t_1]$ называется допустимой, если $x(t_0)=x_0$, $x(t_1)=x_1$. Множество всех допустимых функций обозначается символом $\mathcal{A}$.*

Определение 2.2. *Функция $\hat{x}(t)\in\mathcal{A}$ называется точкой минимума (максимума) задачи (1),(2), если $\forall x(t)\in\mathcal{A}$ выполнено неравенство $J(\hat{x})\le J(x)$ $(J(\hat{x})\ge J(x))$. Точка минимума или максимума называется точкой экстремума.*

Определение 2.3. *Функция $\hat{x}(t)$ называется точкой слабого локального минимума (слабого локального максимума), если $\exists\varepsilon\forall x(t)\in\mathcal{A}$ такого, что $\|\hat{x}-x\|_{C^1[t_0,t_1]}\le\varepsilon$ выполняется неравенство $J(\hat{x})\le J(x)$ $(J(\hat{x}))\ge J(x)$. Точкой слабого локального экстремума называется функция, которая является либо точкой слабого локального минимума, либо точкой слабого локального максимума.*

Наша ближайшая цель-понять, как решать задачи вариационного исчисления вида (1),(2).

2.2. **Необходимое условие экстремума. Уравнение Эйлера.** Справедлива следующая

Теорема 2.1. *Пусть $\hat{x}$-точка локального экстремума задачи (1),(2). Тогда $\hat{x}(t)$ удовлетворяет следующему дифференциальному уравнению:*

$$-\frac{d}{dt}L_{\dot{x}}(t,\hat{x}(t),\dot{\hat{x}}(t))+L_x(t,\hat{x}(t),\dot{\hat{x}}(t))=0, \tag{3}$$

которое называется уравнением Эйлера.

Доказательство. Пусть $h(t) \in C^1[t_0, t_1],\ h(t_0) = h(t_1) = 0,\ \varphi(\lambda) = \int_{t_0}^{t_1} L(t, \hat{x} + \lambda h, \dot{\hat{x}} + \lambda \dot{h})dt$. Тогда

$$0 = \varphi'(\lambda)|_{\lambda=0} = \frac{d}{d\lambda} J(\hat{x} + \lambda h)|_{\lambda=0} =$$

$$= \int_{t_0}^{t_1} L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))\dot{h}(t) + L_x(t, \hat{x}(t), \dot{\hat{x}}(t))h(t)dt \tag{4}$$

Чтобы вывести из (4) равенство (3), нам понадобится следующая лемма, доказательство которой будет дано сразу после окончания доказательства Теоремы 2.1.

Лемма 2.1. *(Дюбуа-Раймон) Пусть* $a(t), b(t) \in C[t_0, t_1]$ *и*

$$\int_{t_0}^{t_1} (a(t)\dot{h}(t) + b(t)h(t))dt = 0 \quad \forall h \in C^1[t_0, t_1] : h(t_0) = h(t_1) = 0. \tag{5}$$

Тогда $a(t) \in C^1[t_0, t_1]$ *и при этом* $-\frac{d}{dt}a(t) + b(t) = 0$.

Закончим доказательство теоремы. Возьмем $a(t) = L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)),\ b(t) = L_x(t, \hat{x}(t), \dot{\hat{x}}(t))$ и, применяя к (4) Лемму Дюбуа-Раймона, получим равенство (3). □

Доказательство. леммы Дюбуа-Раймона. Интегрируя по частям в интеграле из (5), получим, что

$$\int_{t_0}^{t_1} \left(a(t) - \int_{t_0}^{t} b(\tau)d\tau \right) \dot{h}(t)dt = 0, \tag{6}$$

откуда мы хотим вывести, что

$$\varphi(t) \equiv a(t) - \int_{t_0}^{t} b(\tau)d\tau \equiv \text{const}. \tag{7}$$

Предположим, что (7) не верно, то есть существуют такие $t_0 < \tau_0 < \tau_1 < t_1$, что $\varphi(\tau_0) \neq \varphi(\tau_1)$, например, $\varphi(\tau_0) > \varphi(\tau_1)$. В силу непрерывности $\varphi(t)$, существует такое $\delta > 0$, что $t_0 < \tau_0 - \delta < \tau_0 + \delta < \tau_1 - \delta < \tau_1 + \delta < t_1$ и

$$\varphi(\tau_0 + s) > \varphi(\tau_1 + s) \quad \forall s \in [-\delta, \delta]. \tag{8}$$

Пусть

$$\dot{h}(t) = \begin{cases} -(t - \tau_0 + \delta)(t - \tau_0 - \delta), & t \in (\tau_0 - \delta, \tau_0 + \delta), \\ (t - \tau_1 + \delta)(t - \tau_1 - \delta), & t \in (\tau_1 - \delta, \tau_1 + \delta), \\ 0, & t \in [t_0, t_1] \setminus ((\tau_0 - \delta, \tau_0 + \delta) \cup (\tau_1 - \delta, \tau_1 + \delta)). \end{cases}$$

Подставляя в равенство (6) функцию $\dot{h}(t)$, определенную выше и $\varphi(t)$, определенную в (7), получим, что

$$-\int_{\tau_0-\delta}^{\tau_0+\delta} \varphi(t)(t - \tau_0 + \delta)(t - \tau_0 - \delta)dt + \int_{\tau_1-\delta}^{\tau_1+\delta} \varphi(t)(t - \tau_1 + \delta)(t - \tau_1 - \delta)dt =$$

$$= \int_{-\delta}^{\delta} (\varphi(s + \tau_1) - \varphi(s + \tau_0))(s + \delta)(s - \delta)ds = 0$$

При этом первое из равенств получается в результате замены переменных $t - \tau_0 = s$ в первом из интегралов в левой части и замены $t - \tau_1 = s$ во втором из интегралов. Второе же из равенств, которое должно быть справедливо в силу (6) на самом деле не выполняется, так как в силу (8) в интеграле из левой части последнего из равенств при любом $s \in (-\delta, \delta)$ два сомножителя строго отрицательна, а один строго положителен. Полученное противоречие доказывает второе из равенств (7), из которого следует лемма. $\square$

2.3. **Решение простейшего примера.** Рассмотрим следующую простейшую задачу вариационнго исчисления:

$$\int_0^1 \dot{x}^2(t)dt \to \text{extr}, \quad x(0) = 0,\ x(1) = 1$$

Здесь $L(t, x, \dot{x}) = \dot{x}^2$, то есть $-\frac{d}{dt}L_{\dot{x}} + L_x = -2\ddot{x} = 0$, откуда $x(t) = at + b, \hat{x}(t) = t$.

Проверка. Возьмем произвольную функцию $h(t) \in C^1[0,1], h(0) = h(1) = 0$. Тогда

$$\int_0^1 (\dot{\hat{x}} + \dot{h})^2 - (\dot{\hat{x}})^2 dt = \int_0^1 (1 + \dot{h})^2 - 1dt = \int_0^1 2\dot{h}(t) + \dot{h}^2(t)dt = \int_0^1 \dot{h}^2(t)dt \geq 0$$

Ответ: Полученное решение является абсолютным минимумом.

Лекция 3

2.4. **Первый интеграл уравнения Эйлера.** Это понятие нам понадобится при решении задачи о брахистохроне. Напомним, что решение $\hat{x}(t)$ простейшей задачи вариационного исчисления

$$J(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \to \text{extr}, \quad x(t_0) = x_0, \quad x(t_1) = x_1$$

удовлетворяет уравнению Эйлера

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

Определение 2.4. *Функция $\Phi(t, x, \dot{x})$ называется первым интегралом уравнения Эйлера, если для любого решения $\hat{x}(t)$ уравнения Эйлера справедливо равенство $\Phi(t, \hat{x}(t), \dot{\hat{x}}(t) \equiv const$.*

Приведем примеры первых интегралов.

а) Интеграл импульса (когда $L(t, x, \dot{x}) \equiv L(t, \dot{x})$ не зависит от x). Он имеет вид: $\Phi(t, \dot{x}) = L_{\dot{x}}(t, \dot{x})$. Покажем, что Φ- первый интеграл. Действительно, так как $L_x(t, \dot{x}) = 0$, то для любого решения $\hat{x}(t)$ уравнения Эйлера

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

Значит $\Phi(t, \dot{\hat{x}}(t)) = L_{\dot{x}}(t, \dot{\hat{x}}(t)) \equiv \text{const}$.

б) Интеграл энергии (когда $L(t, x, \dot{x}) \equiv L(x, \dot{x})$ не зависит от t). Он имеет вид: $\Phi(x, \dot{x}) = L_{\dot{x}}\dot{x} - L$. Покажем, что Φ- первый интеграл. Действительно, для любого решения $\hat{x}(t)$ уравнения Эйлера, используя обозначение $\hat{L}(t) = L(\hat{x}(t), \dot{\hat{x}}(t))$, получим, что

$$\frac{d}{dt}\Phi(\hat{x}(t), \dot{\hat{x}}(t)) = \frac{d}{dt}(\hat{L}_{\dot{x}}\dot{\hat{x}} - \hat{L}) = \frac{d}{dt}(\hat{L}_{\dot{x}})\dot{\hat{x}} - \hat{L}_x\dot{\hat{x}} - \hat{L}_{\dot{x}}\ddot{\hat{x}} + \hat{L}_{\dot{x}}\ddot{\hat{x}} = \frac{d}{dt}(\hat{L}_{\dot{x}})\dot{\hat{x}} - \hat{L}_x\dot{\hat{x}} = 0,$$

и значит $\Phi(\hat{x}(t), \dot{\hat{x}}(t)) \equiv \text{const}$.

2.5. **Решение задачи о брахистохроне.** Напомним, что задача о брахистохроне имеет следующий вид:

$$J(x) = \int_0^T \frac{\sqrt{1+(\dot{x}(t))^2}}{\sqrt{x(t)}} dt \to \min, \quad x(0) = 0,\ x(T) = x_T. \tag{9}$$

(Эта формулировка отличается от формулировки, полученной при выводе этой задачи тем, что мы переобозначили искомую функцию с $y(x)$ на $x(t)$, область определения этой функции с $[0, x_1]$ на $[0, T]$ и убрали в функционале J множитель $1/\sqrt{2g}$. Это никак не изменило по сути саму задачу.)

Отметим, что подинтегральная функция в минимизируемом функционале имеет вид $L(x, \dot{x}) = \sqrt{\frac{1+\dot{x}^2(t)}{x(t)}}$, то есть она не зависит явно от t. Поэтому у уравнения Эйлера рассматриваемой задачи имеется первый интеграл энергии

$$\Phi(x, \dot{x}) = L_{\dot{x}}\dot{x} - L = \frac{\dot{x}^2}{\sqrt{(1+\dot{x}^2)x}} - \frac{1+\dot{x}^2}{\sqrt{(1+\dot{x}^2)x}} = \frac{-1}{\sqrt{(1+\dot{x}^2)x}}$$

Другими словами, любое решение $\hat{x}(t)$ уравнения Эйлера удовлетворяет соотношению $(1 + \dot{\hat{x}}^2(t))\hat{x}(t) \equiv \mathrm{const} = c$, то есть $\hat{x}(t)$ является решением дифференциального уравнения

$$\dot{x} = \sqrt{\frac{c}{x} - 1} = \frac{dx}{dt} \quad \Leftrightarrow \quad dx = \sqrt{\frac{c}{x} - 1} dt. \tag{10}$$

Сделаем в этом уравнении замену переменных

$$x(\tau) = c\sin^2\frac{\tau}{2} \ \Leftrightarrow\ dx = c\sin\frac{\tau}{2}\cos\frac{\tau}{2}d\tau = \frac{c}{2}sin\tau d\tau$$

Подставляя эти соотношения в (10) получим выражение для dt:

$$\frac{c}{2}\sin\tau d\tau = \sqrt{\frac{1}{\sin^2\frac{\tau}{2}} - 1}dt = \mathrm{ctg}\,\frac{\tau}{2}dt \ \Rightarrow\ dt = \frac{c\sin\tau}{2\,\mathrm{ctg}\,\frac{\tau}{2}}d\tau = c\sin^2\frac{\tau}{2}d\tau = \frac{c}{2}(1-\cos\tau)d\tau$$

Из приведенных выше соотношений, учитывая, что $x = 0$ при $t = 0$, получим следующие выражения, определяющие кривую $(t(\tau), x(\tau))$, которая называется циклоидой:

$$t(\tau) = \frac{c}{2}(\tau - \sin\tau),\ x(\tau) = \frac{c}{2}(1 - \cos\tau)$$

Эта кривая задаёт решение задачи (9) о брахистохроне. При этом константа c определяется из краевого условия $x(T) = x_T$ в задаче (9). Действительно, при τ, растущем от 0 до π, функция

$$\frac{t(\tau)}{x(\tau)} = \frac{\tau - \sin\tau}{1 - \cos\tau}$$

монотонно растет от 0 до ∞. Поэтому, так как $x_T > 0$, существует единственное $\hat{\tau} \in (0, \pi)$ такое, что $T/x_T = (\hat{\tau} - \sin\hat{\tau})/(1 - \cos\hat{\tau})$. Следовательно $c = 2T/(\hat{\tau} - \sin\hat{\tau})$.

Замечание 2.1. *Из формулы для циклоиды следует, что при $\tau \to 0$, функции $t(\tau), x(\tau)$ имеют следующие асимптотики: $t \sim c_1\tau^3, x \sim c_2\tau^2$, где c_1, c_2- положительные константы. Поэтому, записывая циклоиду в окрестности нуля как функцию $x(t)$, получим при $t \to 0$ следующую асимптотику этой функции: $x(t) \sim c_3 t^{2/3}$, где c_3- положительная константа.*

2.6. **Задача Больца.** Рассмотрим ещё один класс задач вариационного исчисления, а именно задачи Больца. При этом рассмотрим случай векторных задач Больца. Эта задача записывается следующим образом:

$$J(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt + l(x(t_0), x(t_1))) \to \text{extr} \tag{11}$$

Мы будем рассматрвать случай, когда искомой функцией является векторнозначная функция $x(t) = (x_1(t), x_2(t), \dots, x_n(t)) \in C^1([t_0, t_1]; \mathbb{R}^n)$. Поэтому относительно заданных функций $L(t, x, \dot{x}), l(x(t_0), x(t_1))$ будем предполагать, что $L \in C^1([t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n), l \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$.

Множеством допустимых функций $\mathcal{A}$ в задаче Больца является все пространство $C^1([t_0, t_1]; \mathbb{R}^n)$

Определение 2.5. *Допустимая функция $\hat{x}(t) \in \mathcal{A}$ называется точкой слабого локального минимума (максимума) задачи (11), если существует такое $\varepsilon > 0$, что для любого $x \in \mathcal{A}$, удовлетворяющего условию $\|\hat{x} - x\|_{C^1([t_0,t_1];\mathbb{R}^n)} \leq \varepsilon$, выполнено неравенство $J(\hat{x}) \leq J(x)$ $(J(\hat{x}) \geq J(x))$.*

Как и раньше, точкой слабого локального экстремума задачи (11) мы будем называть точку слабого локального минимума либо максимума.

Ниже мы будем использовать следующие обозначения: $x = (x_1, \dots, x_n)$,

$$L_x(t, x, \dot{x}) = (\frac{\partial L}{\partial x_i}(t, x, \dot{x}), i = 1, \dots, n), \quad L_{\dot{x}}(t, x, \dot{x}) = (\frac{\partial L}{\partial \dot{x}_i}(t, x, \dot{x}), i = 1, \dots, n),$$

$$\hat{L}(t) = L(t, \hat{x}(t), \dot{\hat{x}}(t)), \quad \hat{L}_x(t) = L_x(t, \hat{x}(t), \dot{\hat{x}}(t)), \quad \hat{L}_{\dot{x}}(t) = L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)).$$

Аналогичные обозначения будут использоваться также и для функции $l(x(t_0), x(t_1)))$ из (11).

Справедливо следующее утверждение о необходимых условиях экстремума для задачи Больца.

Теорема 2.2. *Пусть $\hat{x}(t)$ - точка локального экстремума в задаче Больца (11). Тогда $\hat{x}$ удовлетворяет уравнению Эйлера:*

$$-\frac{d}{dt} L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) + L_x(t, \hat{x}(t), \dot{\hat{x}}(t)) = 0 \tag{12}$$

а также условию трансверсальности:

$$\hat{L}_{\dot{x}}(t_j) = (-1)^j \frac{\partial \hat{l}}{\partial x(t_j)}, j = 0, 1 \tag{13}$$

Доказательство. Пусть $\varphi(\lambda) = J(\hat{x} + \lambda h)$. Тогда

$$\varphi'(\lambda)|_{\lambda=0} = \int_{t_0}^{t_1} (L_{\dot{x}}(t, \hat{x}, \dot{\hat{x}}) \cdot \dot{h}(t) + L_x(t, \hat{x}, \dot{\hat{x}}) \cdot h(t))dt + \hat{l}_{x(t_0)} \cdot h(t_0) + \hat{l}_{x(t_1)} \cdot h(t_1) = 0. \tag{14}$$

Возьмем в этом равенстве такое h, что $h(t_0) = h(t_1) = 0$ и $h(t) = (0, \dots 0, h_i(t), 0, \dots, 0)$. Тогда это равенство принимает вид $\int_{t_0}^{t_1} \frac{\partial \hat{L}(t)}{\partial \dot{x}_i} \dot{h}_i(t) + \frac{\partial \hat{L}}{\partial x_i} h_i(t) dt = 0$ и из него с помощью леммы Дюбуа - Раймона следует, что $\hat{x}$ удовлетворяет уравнению Эйлера:

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

Если теперь в равенстве (14) взять $h_i(t_0) = 1,\ h_i(t_1) = 0,\ h(t) = (0, \dots 0, h_i(t), 0, \dots, 0)$, то оно примет следующий вид:

$$\int_{t_0}^{t_1} \frac{\partial \hat{L}}{\partial \dot{x}_i} \dot{h}_i(t) + \frac{\partial \hat{L}}{\partial x_i} h_i(t) dt + \frac{\partial \hat{l}}{\partial x_i(t_0)} = 0$$

Проинтегрировав по частям в первом слагаемом из интеграла в этом равенстве, получим, что

$$\int_{t_0}^{t_1} \left(-\frac{d}{dt} \frac{\partial \hat{L}}{\partial \dot{x}_i} + \frac{\partial \hat{L}}{\partial x_i} \right) h_i(t) dt - \frac{\partial \hat{L}}{\partial \dot{x}_i}(t_0) + \frac{\partial \hat{l}}{\partial x_i(t_0)} = 0$$

В силу (12) интеграл в полученном равенстве равен нулю, и значит это равенство доказывает соотношение (13) при $j = 0, i = 1, \dots, n$. Если провести аналогичные рассуждения, взяв в (14) $h_i(t_0) = 0,\ h_i(t_1) = 1,\ h(t) = (0, \dots 0, h_i(t), 0, \dots, 0)$, то получим соотношение (13) при $j = 1, i = 1, \dots, n$. $\square$

3. Элементы функционального анализа

Напомним некоторые сведения из функционального анализа, необходимые для построения общей теории вариационного исчисления.

3.1. Линейные нормированное и банахово пространства.

Линейное пространство X называется нормированным, если на X задан функционал $\|\cdot\| : X \to \mathbb{R}$, удовлетворяющий следующим условиям:

$$1) \|x\| \geq 0,\ \forall x \in X \quad \text{и } \|x\| = 0 \ \Leftrightarrow\ x = 0,$$

$$2) \quad \|\alpha x\| = |\alpha| \|x\|, \quad \forall x \in X, \quad \forall \alpha \in \mathbb{R},$$

$$3) \|x_1 + x_2\| \leq \|x_1\| + \|x_2\|,\ \forall x_1, x_2 \in X.$$

Функционал $\|\cdot\|$, удовлетворяющий условиям 1),2),3) называется нормой. Чтобы подчеркнуть, что норма задана именно на пространстве X, мы будем писать $\|\cdot\|_X$.

Всякое нормированное пространство X становится метрическим, если ввести расстояние между любыми двумя элементами $x_1, x_2 \in X$ по формуле

$$\rho(x_1, x_2) = \|x_1 - x_2\|$$

Напомним, что последовательность $x_n \in X$ в нормированном пространстве X называется фундаментальной, если

$$\forall \varepsilon\ \exists N(\varepsilon)\ \forall n > N(\varepsilon), m > N(\varepsilon) \quad \|x_m - x_n\| < \varepsilon$$

Определение 3.1. *Нормированное пространство называется полным, если любая его фундаментальная последовательность имеет предел. Полное нормированное пространство называется банаховым.*

Приведем простейшие примеры. а)Пример банахова пространства: пространство непрерывных функций $C[0,1]$, снабженное нормой $\|x\|_{C[0,1]} = \max_{t \in [0,1]} |x(t)|$. б) Пример не банахова пространства: пространство непрерывных функций $C[0,1]$, снабженное нормой $\|x\|_{L_1(0,1)} = \int_0^1 |x(t)| dt$.

Напомним понятие сопряженного пространства, которое будет использоваться в дальнейшем. Пусть X -нормированное пространство с нормой $\|\cdot\|_X$, и

$l(x)$ -линейный функционал на X. Определим норму функционала l с помощью формулы.

$$\|l\| = \sup_{\|x\|_X = 1} |l(x)| \tag{1}$$

Легко проверяется, что $\|l\|$ удовлетворяет условиям 1),2),3), входящим в определение нормы, и значит является нормой. Очевидно, множество линейных функционалов на X с конечной нормой (1) является линейным пространством, а если снабдить его нормой (1), то и линейным нормированным пространством. Это пространство называется сопряженным к X, и обозначается X^*. Справедлива

Теорема 3.1. *Нормированное пространство X^* (с нормой (1)) является полным.*

Доказательство этой теоремы можно найти в главе 4 книги А.Н.Колмогорова, С.В.Фомина "Элементы теории функций и функционального анализа".

Отметим, что установленный в теореме факт о полноте пространства X^* никак не зависит от полноты пространства X.

Пусть X, Y-линейные нормированные пространства с нормами $\|\cdot\|_X, \|\cdot\|_Y$. Рассмотрим произведение этих пространств $X \times Y = \{(x,y) : x \in X, y \in Y\}$, в котором сложение элементов и их умножение на число определяется по формулам

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \quad \alpha(x,y) = (\alpha x, \alpha y),$$

а норма на произведении пространств задается соотношением $\|(x,y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$. Можно показать, что если X и Y-банаховы пространства, то и пространство $X \times Y$ тоже банахово. Значение функционала $x^* \in X^*(y^* \in Y^*, \Lambda \in (x \times Y)^*)$ на элементе $x \in X(y \in Y, (x,y) \in X \times Y)$ будем обозначать $\langle x^*, x\rangle (\langle y^*, y\rangle, \langle \Lambda, (x,y)\rangle)$. Справедлива

Лемма 3.1. *Любой линейный функционал $\Lambda \in (X \times Y)^*$ можно однозначно представить в виде*

$$\langle \Lambda, (x,y)\rangle = \langle x^*, x\rangle + \langle y^*, y\rangle, \tag{2}$$

где $x^* \in X^*$ и $y^* \in Y^*$.

Доказательство. Определим функционалы x^*, y^* по Λ с помощью соотношений $\langle x^*, x\rangle = \langle \Lambda, (x, 0)\rangle$, $\langle y^*, y\rangle = \langle \Lambda, (0, y)\rangle$. Линейность построенных функционалов очевидна, а их ограниченность следует из оценок:

$$|\langle x^*, x\rangle| \le \|\Lambda\| \|(x, 0)\| = \|\Lambda\| \|x\|,$$
$$|\langle y^*, y\rangle| \le \|\Lambda\| \|(0, y)\| = \|\Lambda\| \|y\|$$

Единственность определения x^*, y^* через Λ легко проверяется. □

Лекция 4

3.2. **Факторпространство банахова пространства.** а) Пусть X-линейное пространство, а L - подпространство пространства X. Пусть $x, x' \in X$. Говорят, что x и x' принадлежат одному классу смежности, порожденному подпространством L (и пишут $x \sim x'$), если существует такой вектор $l \in L$, что $x' = x + l$.

Легко проверить, что принадлежность одному классу смежности $x \sim x'$ является соотношением эквивалентности, то есть удовлетворяет следующим трем свойствам: $a)x \sim x, b)x \sim y \Rightarrow y \sim x, c)x \sim y, y \sim z \Rightarrow x \sim z$.

Пусть πx-класс смежности всех элементов, эквивалентных x, то есть $\pi x = \{y \in X : y \sim x\} = x + L$

Определение 3.2. *Факторпространством X/L называется множество всех классов смежности πx.*

Легко видеть, что X/L является линейным пространством.

б) Пусть теперь X-нормированное линейное пространство, $L \subset X$-замкнутое подпространство. Пусть π - линейное отображение, определяемое по формуле $X \ni x \to \pi x \in X/L$. Зададим функционал

$$\|\pi x\|_{X/L} = \inf_{l \in L} \|x + l\|$$

Лемма 3.2. *Функционал $\|\cdot\|_{X/L}$ является нормой.*

Доказательство. 1) Очевидно, $\|\pi x\|_{X/L} \geq 0$. Пусть $\|\pi x\|_{X/L} = 0$. Тогда в силу определения существует такая последовательность $L \ni l_n$, что $x + l_n \to 0$, и значит $l_n \to -x$, откуда $x \in L$.

$$2)\|\alpha \pi x\|_{X/L} = \inf_{l \in L} \|\alpha(x + l)\| \leq |\alpha| \inf_{l \in L} \|x + l\| = |\alpha| \|\pi x\|_{X/L}$$

$$3)\|\pi x + \pi y\|_{X/L} = \inf_{l \in L} \|x + y + l\| = \inf_{l_1 \in L, l_2 \in L} \|x + l_1 + y + l_2\| \leq$$

$$\inf_{l_1 \in L} \|y + l_1\| + \inf_{l_2 \in L} \|y + l_2\| = \|\pi x\|_{X/L} + \|\pi y\|_{X/L}$$

□

Следствие 3.1. $\forall \xi \in X/L; \exists x \in \xi : \|x\| \leq 2\|\xi\|_{X/L}$

Лемма 3.3. *Если X - банахово пространство, L-замкнутое подпространство в X, то X/L является банаховым пространством.*

Доказательство. Пусть ξ_n- фундаментальная последовательность в X/L, то есть $\forall \varepsilon \exists N(\varepsilon) > 0 : \forall n \geq N(\varepsilon), m \geq 1$ $\|\xi_{n+m} - \xi_n\|_{X/L} < \varepsilon$. Положив $\varepsilon_k = 2^{-k}$, выберем такие $n_k \geq N(2^{-k})$, что $\|\xi_{n_k+m} - \xi_{n_k}\|_{X/L} \leq 2^{-k}, \forall m \geq 1$. Таким образом $\|\xi_{n_2} - \xi_{n_1}\|_{X/L} \leq 2^{-1}$, и значит в силу Следствия 3.1 существуют такие $x_1 \in \xi_{n_1}, x_2 \in \xi_{n_2}$, что $\|x_2 - x_1\| \leq 1$. Далее $\|\xi_{n_3} - \xi_2\|_{X/L} \leq 2^{-2}$ и существует $x_3 \in \xi_{n_3}$ такое, что $\|x_3 - x_2\| \leq 2^{-1}, \ldots, \exists x_{n_k} \in \xi_{n_k} : \|x_k - x_{k-1}\| \leq 2^{-k+2}$. Легко видеть, что в силу полноты пространства X последовательность $\{x_k, k = 1, 2, \ldots\}$ сходится: $x_k \to \hat{x}$. Поэтому для класса смежности $\xi_{\hat{x}} = \pi \hat{x}$ получим, что $\|\xi_{n_k} - \xi_{\hat{x}}\|_{X/L} = \min_{l \in L} \|x_k - \hat{x} + l\| \leq \|x_k - \hat{x}\| \to 0$ при $k \to \infty$. □

3.3. **Базовые теоремы и их следствия.** Сформулируем три важные теоремы, доказательства которых можно найти в учебнике А.Н.Колмогорова, С.В.Фомина Функциональный анализ, и докажем важные для нас следствия из этих теорем. Первая из упомянутых теорем - это теорема о продолжении:

Теорема 3.2. *(Хан-Банах) Пусть X - линейное нормированное пространство и L - его замкнутое подпространство. Пусть - l линейный функционал, заданный на L и имеющий на L конечную норму:*

$$\|l\| = \sup_{x \in L, \|x\|=1} l(x) < \infty.$$

Тогда существует линейный ограниченный функционал Λ , заданный на X, такой что $\Lambda(x) = l(x)$ $\forall x \in L$ и $\forall x \in X$ $|\Lambda(x)| \leq \|l\|\|x\|$

Следствие 3.2. *Пусть X - нормированное пространство и $x_0 \in X$, $x_0 \neq 0$. Тогда найдется функционал $\Lambda \in X^*$, такой, что $\|\Lambda\| = 1$ и $\langle \Lambda, x_0 \rangle = \|x_0\|$.*

Доказательство. Определим одномерное линейное подпространство $L \subset X$ по формуле $L = \{x = \lambda x_0, \lambda \in \mathbb{R}\}$. На L зададим линейный функционал l формулой $\langle l, \lambda x_0 \rangle = \lambda \|x_0\|$. Очевидно $\|l\| = 1$. Поэтому утверждение следствия сразу вытекает из теоремы Хана-Банаха. □

Вторая из теорем, упомянутых выше -это теорема об обратном операторе:

Теорема 3.3. *(Банах) Пусть X, Y - банаховы пространства, $A : X \to Y$ - линейный ограниченный оператор, у которого $KerA = 0$, $ImA = Y$. В силу этих условий, очевидно, существует линейный оператор $A^{-1}Y \to X$, обратный оператору A. Утверждается, что оператор A^{-1} непрерывен.*

Для полной ясности формулировки этой теоремы напомним, что $\mathrm{Ker}A = \{x \in X : Ax = 0\}$, $\mathrm{Im}A = \{y \in Y : \exists x \in X : Ax = y\}$, а также, что оператор $A^{-1} : X \to Y$ называется обратным оператору $A : X \to Y$, если $A^{-1}Ax = x \; \forall x \in X$ и $AA^{-1}y = y \; \forall y \in Y$.

Напомним также (это нам понадобится в дальнейшем), что оператор $A : X \to Y$ называется мономорфизмом, если $\mathrm{Ker}A = 0$ и называется эпиморфизмом, если $\mathrm{Im}A = Y$.

Докажем следующую *лемму о правом обратном отображении*, являющуюся следствием теоремы 3.3 Банаха.

Лемма 3.4. *Пусть X, Y - банаховы пространства, $A : X \to Y$ - линейный ограниченный оператор, причем $ImA = Y$. Тогда существует такое отображение $M : Y \to X$ (вообще говоря разрывное и нелинейное), что а) $AM(y) = y \; \forall y \in Y$, б) $\|M(y)\| \leq c\|y\|$ (то есть M -непрерывно в нуле).*

Доказательство. Так как оператор A непрерывен, то подпространство $\mathrm{Ker}A$ замкнуто, и в силу Леммы 3.3 факторпространство $X/\mathrm{Ker}A$ банахово. Определим оператор $\tilde{A} : X/\mathrm{Ker}A \to Y$ с помощью формулы $\tilde{A}\xi_x = Ax$, где $\xi_x = x + \mathrm{Ker}A$-класс смежности, порожденный подпространством $\mathrm{Ker}A$ и содержащий элемент x. Это определение корректно, так как если $x_1 = x + a$, где $a \in \mathrm{Ker}A$, то $\tilde{A}\xi_x = \tilde{A}\xi_{x_1}$.

Отметим, что непосредственно из определения следует, что оператор $\tilde{A} : X/\mathrm{Ker}A \to Y$ является мономорфизмом, а то, что он эпиморфизм следует из эпиморфности оператора A. Поэтому согласно Теореме 3.3 Банаха существует обратный оператор $\tilde{A}^{-1}Y \to X/\mathrm{Ker}A$. Пусть $\xi = \tilde{A}_{-1}y \in X/\mathrm{Ker}A$. Тогда согласно Следствию 3.1 существует такое $x = x(\xi)$, что $\xi = x + \mathrm{Ker}A$, причем $\|x\|_X \leq 2\|\xi\|_{X/\mathrm{Ker}A}$. Описанное (неоднозначное) отображение $\xi \to x(\xi)$ обозначим через π^{-1} и положим $M = \pi^{-1}\tilde{A}^{-1}$ то есть $M(y) = x(\xi)$. Отсюда вытекает равенство $A \cdot M(y) = Ax(\xi) = \tilde{A}\xi = \tilde{A}\tilde{A}^{-1}(y) = y$, а также важная оценка $\|M(y)\| = \|x(\xi)\| \leq 2\xi\|_{X/\mathrm{Ker}A} \leq 2\|\tilde{A}^{-1}\|\|y\|$. □

Наконец, сформулируем третью теорему. Напомним, что подмножество A линейного пространства X называется выпуклым, если для любых $x_1, x_2 \in A$ и $\alpha \in (0, 1)$ справедливо включение $\alpha x_1 + (1 - \alpha)x_2 \in A$.

Теорема 3.4. *(Вторая теорема отделимости) Пусть A - выпуклое замкнутое подмножество нормированного пространства X и $x_0 \in X \setminus A$. Существует такой функционал $x^* \in X^*$, что*

$$\sup_{x \in A} \langle x^*, x \rangle < \langle x^*, x_0 \rangle$$

Говорят, что функционал x^* разделяет множество A и элемент x_0. Очевидно, что $x^* \neq 0$.

Пусть A - подмножество линейного пространства X. Аннулятором A называется множество

$$A^{\perp} = \{x^* \in X^* : \langle x^*, x \rangle = 0 \; \forall x \in A\}$$

Лемма 3.5. *(О нетривиальности аннулятора) Пусть X - линейное нормированное пространство, L - замкнутое подпространство X, причем $X \setminus L \neq \emptyset$. Тогда $L^{\perp} \neq 0$.*

Доказательство. Пусть $x_0 \in X \setminus L$. Согласно второй теореме отделимости (т.е. Теореме 3.4) существует такой функционал $x^* \in X^*$, что

$$\sup_{x \in L} \langle x^*, x \rangle < \langle x^*, x_0 \rangle \tag{3}$$

Покажем, что левая часть в (3) равна нулю. Пусть нет. Тогда для некоторого $x_1 \in L$ $\langle x^*, x_1 \rangle > 0$ (это выражение не может быть отрицательным, так как супремум в (3) берется по линейному пространству). Более того, по той же причине можно левую часть в (3) заменить на $\lim_{\lambda \to \infty} \langle x^*, \lambda x_1 \rangle$. Но полученная величина равна бесконечности, в то время как правая часть (3) конечна. В силу полученного противоречия левая часть (3) действительно 0. Но неравенство (3) строгое, и поэтому $x^* \neq 0$. □

Лемма 3.6. *(О замкнутости образа) Пусть X, Y, Z -банаховы пространства, $A : X \to Y$, $B : X \to Z$ - линейные ограниченные операторы, у которых замкнуты ImA и $BKerA$. Пусть $C : X \to Y \times Z$ - оператор, который определяется по формуле $Cx = (Ax, Bx)$. Тогда образ этого оператора (ImC) замкнут.*

Доказательство. Пусть $(y_0, z_0) \in \overline{\mathrm{Im}C}$, то есть существует такая последовательность $\{x_n : n = 1, 2, \dots\}$, что $Ax_n \to y_0$, $Bx_n \to z_0$ при $n \to \infty$. Из замкнутости множества $\mathrm{Im}A$ следует, что $y_0 \in \mathrm{Im}A$. В силу леммы 3.4 (о правом обратном отображении) у оператора $A : X \to \mathrm{Im}A$ существует такое отображение $M : \mathrm{Im}A \to X$, что $AM(y) = y$ для любого $y \in \mathrm{Im}A$, и при этом $\|M(y)\| \leq c_0 \|y\|$. В силу последнего неравенства, если $\xi_n = M(Ax_n - y_0)$, то $\|\xi_n\| \leq c_0 \|Ax_n - y_0\| \to 0$.

Из последних двух соотношений следует, что

$$A(x_n - \xi_n) = Ax_n - Ax_n + y_0 = y_0, \; B(x_n - \xi_n) = Bx_n - B\xi_n \to z_0,$$

т.е. z_0 принадлежит замыканию множества

$$\Sigma = \{\xi = Bx : \; Ax = y_0\}. \tag{4}$$

Это множество является сдвигом подпространства $B\mathrm{Ker}A$. Действительно, если $\xi_1, \xi_2 \in \Sigma$, то существуют такие $x_1, x_2 \in X$, что $\xi_i = Bx_i, Ax_i = y_0, \; i = 1, 2$, откуда $\xi_2 - \xi_1 = B(x_2 - x_1)$, где $A(x_2 - x_1) = 0$, то есть $\xi_2 - \xi_1 \in B\mathrm{Ker}A$. Значит Σ -замкнутое множество, и поэтому $z_0 \in \Sigma$. Следовательно, согласно определению (4) существует такое $\hat{x} \in X$, что $z_0 = B\hat{x}, y_0 = A\hat{x}$. □

Теорема 3.5. *(Об аннуляторе ядра) Пусть X, Y - банаховы пространства, и $A : X \to Y$ - линейных ограниченный оператор, у которого $ImA = Y$. Тогда $(KerA)^{\perp} = ImA^*$.*

Доказательство. (Напомним, что сопряженный оператор $A^* : Y^* \to X^*$ является линейным и ограниченным и определяется по оператору A с помощью формулы $\langle A^*y^*, x\rangle = \langle y^*, Ax\rangle \ \forall x \in X, \ y^* \in Y^*$)

Докажем сначала, что $\mathrm{Im}A^* \subset (\mathrm{Ker}A)^{\perp}$. Пусть $x^* \in \mathrm{Im}A^*$ то есть существует такой $y^* \in Y^*$, что $x^* = A^*y^*$. Пусть $x \in \mathrm{Ker}A$. Тогда $\langle x^*, x\rangle = \langle A^*y^*, x\rangle = \langle y^*, Ax\rangle = 0$.

Докажем теперь, что $(\mathrm{Ker}A)^{\perp} \subset \mathrm{Im}A^*$. Пусть $x^* \in (\mathrm{Ker}A)^{\perp}$, то есть $\langle x^*, x\rangle = 0$ для любого $x \in \mathrm{Ker}A$. Рассмотрим оператор $C : X \to \mathbb{R} \times Y$, определяемый формулой $Cx = (\langle x^*, x\rangle, Ax)$. Так как образ оператора $\mathrm{Ker}A \ni x \to \langle x^*, x\rangle$, очевидно, замкнут, то по Лемме 3.6 о замкнутости образа $\mathrm{Im}C$ замкнут в $\mathbb{R} \times Y$.

Покажем, что $\mathrm{Im}C \subsetneqq \mathbb{R} \times Y$. Предполагая противное, получим, что для любого $(\alpha, 0) \in \mathbb{R} \times Y$ с отличным от нуля α существует такое x, что $\langle x^*, x\rangle = \alpha$ и $Ax = 0$. Но это не верно, так как $x^* \in (\mathrm{Ker}A)^{\perp}$.

По Лемме 3.5 о нетривиальности аннулятора существует такой ненулевой функционал $\Lambda \subset (\mathbb{R} \times Y)^*$, что $0 = \langle \Lambda, (\langle x^*, x\rangle, Ax)\rangle = \lambda_0\langle x^*, x\rangle + \langle y^*, Ax\rangle$. Если предположить, что $\lambda_0 = 0$ мы получим, что $\langle y^*, Ax\rangle = \langle y^*, y\rangle$ для любого $y \in Y$, то есть $y^* = 0$, чего быть не может. Если же $\lambda_0 = 1$, то $\langle x^* + A^*y^*, x\rangle = 0$ для любого $x \in X$, и значит $x^* = -A^*y^* \in \mathrm{Im}A^*$. □

Лекция 5

4. Основы дифференциального исчисления в нормированных пространствах

Приведем основные сведения о дифференциальном исчислении в нормированных пространствах, необходимые для построения общей теории вариационного исчисления. Естественно, они являются прямым обобщением теории дифференциального исчисления в конечномерных линейных пространствах.

4.1. **Определения и простейшие свойства производных.** Начнем с определений. Пусть X, Y - нормированные пространства, $\hat{x} \in X$, $U(\hat{x})$ окрестность $\hat{x}$, то есть открытое подмножество в X, содержащее $\hat{x}$. Примером окрестности является шар $U(\hat{x}) = \{x \in X : \|x - \hat{x}\| < \varepsilon\}$. Рассмотрим отображение $f : U(\hat{x}) \to Y$.

Определение 4.1. *Отображение $f : U(\hat{x}) \to Y$ дифференцируемо по Гато в точке $\hat{x}$, если существует такой линейный ограниченный оператор $\Lambda : X \to Y$, что для любых $h \in X$ и $\lambda \in \mathbb{R}$ таких, что $\hat{x} + \lambda h \in U(\hat{x})$ справедливо равенство*

$$f(\hat{x} + \lambda h) - f(\hat{x}) - \lambda\Lambda h = \overline{o}(\lambda), \text{ где по определению } \frac{\|\overline{o}(\lambda)\|}{|\lambda|} \to 0 \text{ при } \lambda \to 0. \quad (1)$$

При этом оператор Λ называется производной Гато отображения f в точке $\hat{x}$ и обозначается $\Lambda = f'_{\Gamma}(\hat{x})$.

Определение 4.2. *Отображение $f : U(\hat{x}) \to Y$ дифференцируемо по Фреше в точке $\hat{x}$, если существует такой линейный ограниченный оператор $\Lambda : X \to$*

Y, что для любого $h \in X$ такого, что $\hat{x} + h \in U(\hat{x})$ справедливо равенство

$$f(\hat{x} + h) - f(\hat{x}) - \Lambda h = \overline{o}(h), \text{ где } \frac{\|\overline{o}(h)\|}{\|h\|} \to 0 \text{ при } \|h\| \to 0. \qquad (2)$$

При этом оператор Λ называется производной Фреше отображения f в точке $\hat{x}$ и обозначается $\Lambda = f'(\hat{x})$.

Определение 4.3. *Отображение $f : U(\hat{x}) \to Y$ строго дифференцируемо в точке $\hat{x}$, если существует такой линейный ограниченный оператор $\Lambda : X \to Y$, что $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x_1, x_2 : \|x_1 - \hat{x}\| < \delta, \|x_2 - \hat{x}\| < \delta$ справедливо равенство*

$$\|f(x_1) - f(x_2) - \Lambda(x_1 - x_2)\| < \varepsilon \|x_1 - x_2\|. \qquad (3)$$

Докажем следующее утверждение.

Лемма 4.1. *Производная Гато определяется однозначно.*

Доказательство. Предположим, что у отображения $f : U(\hat{x}) \to Y$ в точке $\hat{x}$ имеется две производных Гато Λ_1 и Λ_2, каждая из которых удовлетворяет своему равенству (1). Вычитая из равенства для Λ_1 равенство для Λ_2, получим, что $-\lambda(\Lambda_1 - \Lambda_2)h = \overline{o}_1(\lambda) - \overline{o}_2(\lambda)$. Пусть для определенности в этом соотношении $\lambda > 0$. Поделив его на $|\lambda|$, получим, учитывая определение $\overline{o}(\lambda)$ из (1), что $(\Lambda_2 - \Lambda_1)h = \frac{\overline{o}_1(\lambda) - \overline{o}_2(\lambda)}{|\lambda|} \to 0$ при $|\lambda| \to 0$. Следовательно, $(\Lambda_1 - \Lambda_2)h = 0$ при любом $h \in X$ и значит $\Lambda_1 = \Lambda_2$. □

Ниже из этой леммы будет получена однозначность определения производной Фреше и производной в случае строгой дифференцируемости.

Выясним, как связаны понятия производных, введенные выше.

Предложение 4.1. *Пусть отображение $f : U(\hat{x}) \to Y$ дифференцируемо по Фреше в точке $\hat{x}$. Тогда оно также дифференцируемо и по Гато в этой точке, причем его производные Фреше и Гато совпадают: $f'(\hat{x}) = f'_\Gamma(\hat{x})$. Следовательно производная Фреше $f'(\hat{x})$ определяется однозначно.*

Доказательство. Предположение о дифференцируемости по Фреше означает справедливость для f Определения 4.2. Взяв в этом определении $h \in X$ специального вида, а именно $h = \lambda h_0$, где $h_0 \in X$-фиксированный вектор, а $\lambda \in \mathbb{R}$, получим в точности Определение 4.1 дифференцируемости по Гато, что доказывает дифференцируемость по Гато и совпадение производных Гато и Фреше отображения f в точке $\hat{x}$. Отсюда, учитывая Лемму 4.1, получаем однозначность определения производной Фреше $f'(\hat{x})$. □

Предложение 4.2. *Пусть отображение $f : U(\hat{x}) \to Y$ строго дифференцируемо в точке $\hat{x}$. Тогда оно дифференцируемо и по Фреше в $\hat{x}$, причем ее строгая производная $f'(\hat{x})$ совпадает с производной Фреше. Отсюда следует однозначность определения строгой производной $f'(\hat{x})$.*

Доказательство. Заменив соотношение (2) на

$$f(\hat{x} + h) - f(\hat{x}) - \Lambda h = \alpha(h), \text{ где } \alpha(h) \to 0 \text{ при } \|h\| \to 0,$$

получим определение, полностью эквивалентное Определению 4.2. Переписав его на $\varepsilon\delta$ -языке, получим следующую форму определения дифференцируемости по Фреше отображения f в точке $\hat{x}$: для любого $\varepsilon > 0$ найдется $\delta > 0$, при котором для всех h, таких, что $\|h\| < \delta$, справедливо неравенство

$$\|f(\hat{x} + h) - f(\hat{x}) - \Lambda h\| \le \varepsilon \|h\| \qquad (4)$$

Положив теперь в Определении 4.3 $x_2 = \hat{x}$ и $x_1 = \hat{x} + h$, получим (4), так что строго дифференцируемая функция дифференцируема и по Фреше и $\Lambda = f'(\hat{x})$. Наконец, из однозначности определения производной Фреше, доказанной в Предложении 4.2, следует однозначность определения строгой производной. $\square$

Таким образом, в Предложениях 4.1, 4.2 доказано, что между определениями 4.1, 4.2, 4.3 имеют место следующие импликации:

$$\text{Опр. } 4.1 \Leftarrow \text{Опр. } 4.2 \Leftarrow \text{Опр. } 4.3$$

Покажем что противоположные импликации не верны.

Предложение 4.3. *Не верны следующие импликации:*

$$1)\,\textit{Опр. } 4.1 \not\Rightarrow \textit{Опр. } 4.2, \quad 2)\,\textit{Опр. } 4.2 \not\Rightarrow \textit{Опр. } 4.3$$

Доказательство. Отсутствие импликаций можно доказать с помощью контрпримеров. Для доказательства утверждения 1), т.е., что из дифференцируемости по Гато не следует дифференцируемость по Фреше, рассмотрим отображение $f : \mathbb{R}^2 \to \mathbb{R}$, заданное формулой

$$f(x_1, x_2) = \begin{cases} 1, & \text{при } x_1 = x_2^2,\ x_2 > 0, \\ 0, & \text{в остальных точках.} \end{cases}$$

Очевидно, это отображение дифференцируемо по Гато в точке (0,0) и $f'(0,0) = (0,0)$. При этом оно не дифференцируемо в этой точке по Фреше, так как в этой точке оно разрывно, хотя из Определения 4.2 легко следует, что дифференцируемость по Фреше отображения в некоторой точке влечет его непрерывность в этой точке.

Для доказательства утверждения 2), рассмотрим отображение $g : \mathbb{R} \to \mathbb{R}$, заданное формулой

$$g(x) = \begin{cases} x^2, & \text{при } x \text{ рациональном,} \\ 0, & \text{при } x \text{ иррациональном.} \end{cases}$$

Эта функция, очевидно, дифференцируема по Фреше в нуле и $g'(0) = 0$, но не строго дифференцируема, потому что неравенство (3) не выполняется, если взять в нем x_1 рациональным и $0 < x_1 < \delta$, а x_2 иррациональным, $0 < x_2 < \delta$ и x_2 стремящемся к x_1. $\square$

4.2. **Некоторые теоремы дифференциального исчисления.**

Теорема 4.1. *(о суперпозиции дифференцируемых отображений) Пусть X, Y, Z - банаховы пространства, $\hat{x} \in X$, $\hat{y} \in Y$ и $U := U(\hat{x}) \subset X$, $V := V(\hat{y}) \subset Y$ - окрестности точек $\hat{x}, \hat{y}$. Пусть $\varphi : U \to V$ и $\psi : V \to Z$ - отображения, а $f : U \to Z$ - их суперпозиция, то есть $f(x) = \psi(\varphi(x)) := \psi \circ \varphi$, причем $\hat{y} = \varphi(\hat{x})$. Если отображение Ψ дифференцируемо по Фреше в точке $\hat{y}$, а отображение φ дифференцируемо по Фреше (по Гато) в точке $\hat{x}$, то f дифференцируемо по Фреше (по Гато) в точке $\hat{x}$ и*

$$f'(\hat{x}) = \psi'(\hat{y})\varphi'(\hat{x}) = \psi'(\varphi(\hat{x}))\varphi'(\hat{x}). \tag{6}$$

Если же отображения ψ, φ строго дифференцируемы соответственно в точках $\hat{y}, \hat{x}$, то f строго дифференцируемо в $\hat{x}$ и справедливо соотношение (6).

Эта теорема является одной из базовых теорем дифференциального исчисления, и ее конечномерный аналог доказывается в курсе математического анализа. Обобщение этого доказательства на бесконечномерный случай является прямолинейным и не содержит новых подходов. Поэтому в лекциях оно не приводится. Тем не менее я советую его прочитать. Найти доказательство можно, например, в учебнике В.М.Алексеева, В.М.Тихомирова, С.В.Фомина Оптимальное управление.

Рассмотрим теперь обобщение следующей известной теоремы Лагранжа: Пусть $x_1, x_2 \in \mathbb{R}$, $f \in C[x_1, x_2]$ и дифференцируема в каждой точке интервала (x_1, x_2). Тогда существует такое $x_0 \in (x_1, x_2)$, что

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x_0).$$

Теорема 4.2. *(о среднем). Пусть X, Y - нормированные пространства, U - открытое подмножество в X, и $f : U \to Y$ - отображение, дифференцируемое по Гато в любой точке $x \in U$. Пусть $[x_1, x_2] := \{x = (1-\lambda)x_1 + \lambda x_2, \lambda \in [0,1]\} \subset U$. Тогда $\|f(x_2) - f(x_1)\| \le \sup_{x \in [x_1, x_2]} \|f'(x)\| \|x_2 - x_1\|$*

Доказательство. Пусть $y^* \in Y^*$ -некоторый линейный функционал. На отрезке $\lambda \in [0,1]$ рассмотрим функцию $\Phi(\lambda) = \langle y^*, f((1-\lambda)x_1 + \lambda x_2)\rangle$. В силу условий теоремы к этой функции применима теорема Лагранжа, согласно которой существует такое $\hat{\lambda} \in (0,1)$, что $\Phi(1) - \Phi(0) = \Phi'(\hat{\lambda})$. В силу следствия 3.2 Теоремы Хана-Банаха относительно выбранного выше функционала y^* можно предполагать, что $\|y^*\| = 1$ и $\langle y^*, f(x_2) - f(x_1)\rangle = \|f(x_2) - f(x_1)\|$. Принимая все это в расчет, получим: $\|f(x_2) - f(x_1)\| = \langle y^*, f(x_2) - f(x_1)\rangle = \langle y^*, f'((1-\hat{\lambda})x_1 + \hat{\lambda} x_2)(x_2 - x_1)\rangle \le \sup_{x \in [x_1, x_2]} \|f'(x)\| \|x_2 - x_1\|$. □

Пусть X, Y - линейные нормированные пространства. Напомним, что через $\mathcal{L}(X, Y)$ обозначается множество всех линейных ограниченных операторов $\Lambda : X \to Y$, с конечной нормой $\|\Lambda\| = \sup_{\|x\|=1} \|\Lambda x\|$. Очевидно, что $\mathcal{L}(X, Y)$ - линейное нормированное пространство.

Следствие 4.1. *Пусть $f : U \to Y$ - отображение, дифференцируемое по Гато при всех $x \in U \subset X$, а $\Lambda \in \mathcal{L}(X, Y)$. Тогда при $[x_1, x_2] \subset U$ справедливо неравенство*

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

Лекция 6

Пусть X, Y - линейные нормированные пространства, U - открытое подмножество (область) в X.

Определение 4.4. *Отображение $f : U \to Y$ непрерывно дифференцируемо в U по 1) Гато, 2) Фреше, 3) строго, если для любого $x \in U$ $f(x)$ дифференцируемо по 1) Гато, 2) Фреше, 3) строго, и производная $f'(x) : U \to \mathcal{L}(X, Y)$ непрерывна в любой точке $x \in U$.*

Теорема 4.3. *Если отображение $f : U \to Y$ непрерывно дифференцируемо по Гато в U, то f непрерывно дифференцируемо в U также по Фреше и строго.*

Доказательство. Покажем, что отображение $f(x)$ строго дифференцируемо в любой точке $\hat{x} \in U$, то есть, что

$$\forall \varepsilon\ \exists \delta:\ \forall x_1, x_2 \|x_i - \hat{x}\| < \delta, i = 1,2\ \|f(x_2) - f(x_1) - f'(\hat{x})(x_2 - x_1)\| \leq \varepsilon \|x_2 - x_1\| \tag{7}$$

Действительно, зафиксируем $\hat{x}$ и выберем такое δ_0, что шар $B_{\delta_0}(\hat{x}) = \{x \in X : \|x - \hat{x}\| < \delta_0\}$ радиуса δ_0 с центром в $\hat{x}$ принадлежит U. В силу непрерывной дифференцируемости по Гато отображения f в точке $\hat{x} \in U$

$$\forall \varepsilon\ \exists \delta < \delta_0:\ \forall x\ \|x - \hat{x}\| < \delta \Rightarrow \|f'(x) - f'(\hat{x})\| < \varepsilon \tag{8}$$

В силу следствия 4.1 теоремы о среднем, при выполнении включения $[x_1, x_2] \subset U$ справедливо неравенство

$$\|f(x_2) - f(x_1) - f'(\hat{x})(x_2 - x_1)\| \leq \sup_{x \in [x_1, x_2]} \|f'(x) - f'(\hat{x})\| \|x_2 - x_1\| \tag{9}$$

Для любого $\varepsilon > 0$ выберем такое δ, чтобы выполнялось (8). Возьмем любые x_1, x_2, для которых $\|x_i - \hat{x}\| < \delta, i = 1, 2$. Так как $\delta < \delta_0$, то $[x_1, x_2] \subset U$ и значит выполнено неравенство (9). Более того $[x_1, x_2] \subset B_{\delta_0}(\hat{x})$, и значит последнее неравенство из (8) верно для любого $x \in [x_1, x_2]$. Следовательно из (8),(9) следует (7). Так как строгая производная совпадает с производными Фреше и Гато того же отображения, а производная Гато непрерывна по условию, то доказана непрерывность производной Фреше и строгой производной. □

4.3. **Дифференцирование в произведении пространств. Частные производные.** Обобщим понятие частной производной на случай бесконечномерных отображений.

Пусть X, Y, Z - нормированные пространства, $U \subset X \times Y$ - открытое подмножество в прямом произведении пространств $X \times Y$, $f(x, y) : U \to Z$ - отображение.

Определение 4.5. *Отображение $f(x, y)$ имеет частную производную по x в смысле 1) Гато, 2) Фреше, 3) строгой в точке $(\hat{x}, \hat{y}) \in U$, если его сужение $f(x, \hat{y}) : U|_{y=\hat{y}} \subset X \to Z$ имеет производную в точке $\hat{x}$ по 1) Гато, 2) Фреше, 3) строгой. Эта производная обозначается символом $f'_x(\hat{x}, \hat{y})$ или $\frac{\partial f}{\partial x}$.*

Аналогичным образом определяется частная производная по y, то есть $f'_y(\hat{x}, \hat{y}) = \frac{\partial f}{\partial y}$.

Теорема 4.4. *(О полном дифференциале) Пусть X, Y, Z - нормированные пространства, U - окрестность в $X \times Y$, $f : U \to Z$ - отображение, имеющее в каждой точке $(x, y) \in U$ частные производные Гато $f'_x(\hat{x}, \hat{y})$ и $f'_y(\hat{x}, \hat{y})$.*

Если отображения $(x, y) \to f'_x(\hat{x}, \hat{y})$ и $(x, y) \to f'_y(\hat{x}, \hat{y})$ непрерывны в точке $(\hat{x}, \hat{y}) \in U$ в операторной топологии, то f строго дифференцируемо в той же точке, и при этом

$$f'(\hat{x}, \hat{y})[\xi, \eta] = f'_x(\hat{x}, \hat{y})\xi + f'_y(\hat{x}, \hat{y})\eta$$

Доказательство. Взяв произвольное $\varepsilon > 0$, выберем столь малое $\delta > 0$, что "прямоугольная"окрестность

$$V = B_\delta(\hat{x}) \times B_\delta(\hat{y}) = \{(x, y) : \|x - \hat{x}\| < \delta,\ \|y - \hat{y}\| < \delta\}$$

точки $(\hat{x}, \hat{y})$ содержалась в U, и в ней выполнялись неравенства

$$\|f'_x(x, y) - f'_x(\hat{x}, \hat{y})\| < \varepsilon, \quad \|f'_y(x, y) - f'_y(\hat{x}, \hat{y})\| < \varepsilon. \tag{9}$$

Теперь имеем:

$$\Delta = f(x_1,y_1) - f(x_2,y_2) - f'_x(\hat{x},\hat{y})(x_1-x_2) - f'_y(\hat{x},\hat{y})(y_1-y_2) =$$

$$= [f(x_1,y_1) - f(x_2,y_1) - f'_x(\hat{x},\hat{y})(x_1-x_2)] + [f(x_2,y_1) - f(x_2,y_2) - f'_y(\hat{x},\hat{y})(y_1-y_2)]$$

Очевидно, что если точки (x_1,y_1), (x_2,y_2) лежат в V, то и $(x_2,y_1) \in V$. Более того, оба отрезка $[(x_1,y_1),(x_2,y_1)]$ и $[(x_2,y_1),(x_2,y_2)]$ содержатся в $V \subset U$. Поэтому функции $x \to f(x,y_1)$ и $y \to f(x_2,y)$ дифференцируемы по Гато: первая имеет производную f'_x на $[x_1,x_2]$, вторая f'_y на y_1,y_2. Применяя Следствие 4.1 теоремы о среднем с соответствующим Λ, получаем в силу (9)

$$(x_1,y_1) \in V,\ (x_2,y_2) \in V \Rightarrow |\Delta| \le \sup_{\xi\in[x_1,x_2]} \|f'_x(\xi,y_1) - f'_x(\hat{x},\hat{y})\| \|x_1-x_2\| +$$

$$+ \sup_{\eta\in[y_1,y_2]} \|f'_y(x_2,\eta) - f'_y(\hat{x},\hat{y})\|_1 \|y_1-y_2\| \le \varepsilon\|x_1-x_2\| + \varepsilon\|y_1-y_2\|.$$

□

Следствие 4.2. *Для того, чтобы $f \in C^1(U)$, необходимо и достаточно, чтобы в U частные производные f'_x и f'_y были непрерывны.*

5. Дифференцируемость некоторых конкретных отображений

5.1. **Оператор Немыцкого.** Пусть U - открытое множество в $\mathbb{R}\times\mathbb{R}^n$, и пусть функция $f(t,x) : U \to \mathbb{R}^m$ и ее частная производная $f'_x(t,x)$ непрерывны в U. (Здесь $f'_x(t,x)$ - матрица Якоби с элементами $\frac{\partial f_j}{\partial x_i}(t,x)$, где $i = 1,\dots,n$ - пробегает строку, а $j = 1,\dots,m$ -пробегает столбец.) Зададим отображение

$$\mathcal{N} : W \to C([\hat{t}_0,\hat{t}_1],\mathbb{R}^m), \tag{1}$$

определенное на множестве

$$W = \{x(\cdot) \in C([\hat{t}_0,\hat{t}_1],\mathbb{R}^n) : (t,x(t)) \in U,\ \text{при любом}\ \hat{t}_0 \le t \le \hat{t}_1\} \tag{2}$$

с помощью равенства

$$\mathcal{N}(x(\cdot))(t) \equiv f(t,x(t)),\ \hat{t}_0 \le t \le \hat{t}_1 \tag{3}$$

Это отображение называется оператором Немыцкого.

Теорема 5.1. *При сделанных предположениях оператор Немыцкого (1)-(3) непрерывен и строго дифференцируем в любой точке $\hat{x} \in W$. При этом его производная определяется равенством $(\mathcal{N}'(\hat{x})h)(t) = f'_x(t,\hat{x}(t))h(t)$.*

Доказательство. Пусть $\hat{x} \in W$. Тогда при достаточно малом $\hat{\delta} > 0$ множество $K = \{(t,x) : t \in [\hat{t}_0,\hat{t}_1]$ и $|x - \hat{x}(t)| \le \hat{\delta}\}$ является компактом в $\mathbb{R}^{n+1}$ и компактно вложено в U. Поэтому силу условий теоремы, отображения $f(t,x) : K \to \mathbb{R}^m$ и $f'_x(t,x) : K \to \mathbb{R}^{mn}$ равномерно непрерывны на K, то есть

$\forall\varepsilon\exists\delta : \forall(t_1,x_1),(t_2,x_2) \in K$ и $|t_1-t_2| < \delta, \|x_1-x_2\| < \delta\ \ \|f(t_1,x_1)-f(t_2,x_2)\| \le \varepsilon.$

$\forall\varepsilon_1\exists\delta_1 : \forall(t_1,x_1),(t_2,x_2) \in K, |t_1-t_2| < \delta_1, \|x_1-x_2\| < \delta_1 \|f(t_1,x_1)-f(t_2,x_2)\| \le \varepsilon_1.$

Претендентом на то, чтобы быть производной оператора $\mathcal{N}$ в точке $\hat{x}$ является следующий оператор:

$$\lim_{\lambda\to 0} \frac{f(t,\hat{x}(t)+\lambda h(t)) - f(t,\hat{x}(t))}{\lambda} = f'_x(t,\hat{x}(t))h(t) = (\mathcal{N}'(\hat{x})h)(t).$$

Очевидно, что оператор $h(\cdot) \to (\mathcal{N}'(\hat{x})h)(\cdot)$ линеен по $h(\cdot)$. Поэтому, чтобы доказать, что этот оператор является производной Гато отображения $\mathcal{N}(\hat{x})$, достаточно доказать, что

$$\sup_{t\in[\hat{t}_0,\hat{t}_1]} \|f(t,\hat{x}(t)+\lambda h(t)) - f(t,\hat{x}(t)) - \lambda f'_x(t,\hat{x}(t))h(t)\| = \overline{o}(\lambda), \tag{4}$$

где $\overline{o}(\lambda)$ означает о-малую величину от λ. Чтобы доказать справедливость этого соотношения, возьмем его левую часть, отбросив в нем супремум по t и знаки норм, и зафиксировав $t \in [0,1]$. К полученному выражению применим Следствие 4.1 теоремы о среднем. В результате получим первое из неравенств в

$$\|f(t,\hat{x}(t)+\lambda h(t)) - f(t,\hat{x}(t)) - \lambda f'_x(t,\hat{x}(t))h(t)\| \le$$
$$\le \sup_{\theta\in[0,1]} \|f'_x(t,\hat{x}(t)+\theta\lambda h(t))) - f'_x(t,\hat{x}(t))\|\|\lambda h(t)\| \le \varepsilon\|\lambda h(t)\|. \tag{5}$$

Чтобы получить второе из неравенств (5), по любому выбранному $\varepsilon > 0$ надо взять λ, удовлетворяющее неравенству $0 < \lambda < \lambda_0$, где λ_0 достаточно мало. Следует отметить, что λ_0 можно выбрать не зависящим от $t \in [\hat{t}_0, \hat{t}_1]$, если функции $\hat{x}(t), h(t)$ фиксированы, так как все точки $(t,\hat{x}(t)), (t,\hat{x}(t)+\theta\lambda h(t))$ принадлежат компакту K, введенному в начале доказательства теоремы. Поэтому взяв верхнюю грань по $t \in [\hat{t}_0, \hat{t}_1]$, во всех частях неравенства (5), докажем справедливость соотношения (4) и значит дифференцируемость по Гато отображения $\mathcal{N}(\hat{x})$.

Непрерывная дифференцируемость по Гато следует из оценки

$$\|\mathcal{N}'(\hat{x}) - \mathcal{N}'(x)\| = \sup_{\|h\|_{C([t_0,t_1],\mathbb{R}^n)}=1} \sup_{t\in[t_0,t_1]} \|(f'_x(t,\hat{x}(t)) - f'_x(t,x(t)))h(t)\| \le$$
$$\le \sup_{t\in[t_0,t_1]} \|(f'_x(t,\hat{x}(t)) - f'_x(t,x(t)))\| \le \varepsilon \quad \text{при} \quad \|\hat{x}-x\|_{C([t_0,t_1],\mathbb{R}^n)} < \delta,$$

причем оценка посредством ε при достаточно малых δ опять доказывается с помощью равномерной непрерывности f'_x на компакте K. В силу Теоремы 4.3, из непрерывной дифференцируемости по Гато отображения $\mathcal{N}(\hat{x})$ следует его непрерывная дифференцируемость по Фреше. $\square$

Лекция 7

5.2. **Оператор дифференциальной связи.** Начнем с незначительного обобщения оператора, рассмотренного в предыдущем пункте. Пусть U - открытое множество в $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^r$ и пусть функция $f(t,x,u) : U \to \mathbb{R}^m$ и ее частные производные f'_x, f'_u непрерывны в U. Отображение $\mathcal{N} : W \to C([t_0,t_1],\mathbb{R}^m)$, определенное на множестве

$$W = \{(x(\cdot),u(\cdot)) : (t,x(t),u(t)) \in U, t_0 \le t \le t_1\} \subset C^1([t_0,t_1],\mathbb{R}^n)\times C([t_0,t_1],\mathbb{R}^r) \tag{6}$$

равенством

$$\mathcal{N}(x(\cdot),u(\cdot))(t) \equiv f(t,x(t),u(t)) \tag{7}$$

мы также будем называть оператором Немыцкого.

Предложение 5.1. *Оператор Немыцкого (7) непрерывно дифференцируем на множестве (6) и при этом*

$$\mathcal{N}'(\hat{x}(\cdot),\hat{u}(\cdot))[h(\cdot),v(\cdot)](t) = \hat{f}_x(t)h(t) + \hat{f}_u(t)v(t), \tag{8}$$

где

$$\hat{f}_x(t) \Leftrightarrow \left(\frac{\partial f_i(t,\hat{x}(t),\hat{u}(t))}{\partial x_j}, i=1,\dots,m,\ j=1,\dots,n\right), \tag{9}$$

$$\hat{f}_u(t) \Leftrightarrow \left(\frac{\partial f_i(t,\hat{x}(t),\hat{u}(t))}{\partial u_k}, i=1,\dots,m,\ k=1,\dots,r\right), \tag{10}$$

Доказательство. Как и в предыдущем случае, у отображения $\mathcal{N}$ существуют частные производные

$$\mathcal{N}_x(\hat{x}(\cdot),\hat{u}(\cdot))[h(\cdot)](t) = \hat{f}_x(t)h(t)$$

$$\mathcal{N}_u(\hat{x}(\cdot),\hat{u}(\cdot))[v(\cdot)](t) = \hat{f}_u(t)v(t)$$

и отображения

$$\mathcal{N}_x : W \to \mathcal{L}(C^1([t_0,t_1],\mathbb{R}^n), C([t_0,t_1],\mathbb{R}^m))$$

$$\mathcal{N}_u : W \to \mathcal{L}(C([t_0,t_1],\mathbb{R}^r), C([t_0,t_1],\mathbb{R}^m))$$

непрерывны в W. Остается сослаться на теорему 4.4 о полном дифференциале. □

Пусть U и W будут те же, что и в предыдущем примере. Рассмотрим отображение $\varphi(t,x,u) :\to \mathbb{R}^n$, непрерывное вместе с частными производными φ'_x, φ'_u в U.

Оператором дифференциальной связи называется отображение $\Phi : W \to C([t_0,t_1],\mathbb{R}^n)$, определяемое равенством

$$\Phi(x(\cdot),u(\cdot))(t) = \dot{x}(t) - \varphi(t,x(t),u(t)), \tag{11}$$

Предложение 5.2. *Оператор дифференциальной связи (11) непрерывно дифференцируем на множестве (6) и при этом*

$$\Phi'(\hat{x}(\cdot),\hat{u}(\cdot))[h(\cdot),v(\cdot)](t) = \dot{h}(t) - \hat{\varphi}_x(t)h(t) - \hat{\varphi}_u(t)v(t), \tag{12}$$

где

$$\hat{\varphi}_x(t) \Leftrightarrow \left(\frac{\partial \varphi_i(t,\hat{x}(t),\hat{u}(t))}{\partial x_j},\ i,j=1,\dots,n\right), \tag{13}$$

$$\hat{\varphi}_u(t) \Leftrightarrow \left(\frac{\partial \varphi_i(t,\hat{x}(t),\hat{u}(t))}{\partial u_k},\ i=1,\dots,n\ k=1,\dots,r\right). \tag{14}$$

Это утверждение легко следует из Предложения 5.1.

5.3. **Интегральный функционал.** Пусть U-открытое множество в $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n$, и функция $f(t,x,\dot{x}) : U \to \mathbb{R}^m$ и ее частные производные $f_x, f_{\dot{x}}$ непрерывны в U. Отображение $\mathcal{J} : W \to \mathbb{R}^m$ зададим на множестве

$$W = \{x(\cdot) \in C^1([t_0,t_1],\mathbb{R}^n) : (t,x(t),\dot{x}(t)) \in U, t_0 \le t \le t_1\} \tag{15}$$

равенством

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} f(t,x(t),\dot{x}(t))dt. \tag{16}$$

Предложение 5.3. *Интегральное отображение (16) непрерывно дифференцируемо на множестве (15) и при этом*

$$\mathcal{J}'(\hat{x}(\cdot))[h(\cdot)] = \int_{t_0}^{t_1} \{\hat{f}_x(t)h(t) + \hat{f}_{\dot{x}}\dot{h}(t)\}dt, \tag{17}$$

где

$$\hat{f}_x(t) \Leftrightarrow \left(\frac{\partial f_i}{\partial x_j}(t, \hat{x}(t), \dot{\hat{x}}(t)),\ i = 1, \ldots, m, j = 1, \ldots, n\right), \tag{18}$$

$$\hat{f}_{\dot{x}}(t) \Leftrightarrow \left(\frac{\partial f_i}{\partial \dot{x}_j}(t, \hat{x}(t), \dot{\hat{x}}(t)),\ i = 1, \ldots, m, j = 1, \ldots, n\right). \tag{19}$$

Доказательство. Представим отображение (16) в виде суперпозиции

$$\mathcal{J} = I \circ \mathcal{N} \circ D,$$

где

$$D : C^1([t_0, t_1], \mathbb{R}^n) \to C^1([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^n),\ D(x(\cdot)) \to (x(\cdot), \dot{x}(\cdot))$$

- линейный оператор,

$$\mathcal{N} : U \to C([t_0, t_1], \mathbb{R}^m)$$

- оператор Немыцкого, определяемый равенством (7) (при $r = n$), и

$$I : C([t_0, t_1], \mathbb{R}^m) \to \mathbb{R}^m, \quad I(g(\cdot)) = \int_{t_0}^{t_1} g(t)dt$$

- линейный оператор интегрирования. По теореме о суперпозиции

$$\mathcal{J}'(\hat{x}(\cdot)) = I \circ \mathcal{N}'(D\hat{x}(\cdot)) \circ D,$$

значит

$$\mathcal{J}'(\hat{x}(\cdot))[h(\cdot)] = I\{\mathcal{N}'(D\hat{x}(\cdot))[Dh(\cdot)]\} = \int_{t_0}^{t_1} \{\hat{f}_x(t)h(t) + \hat{f}_{\dot{x}}(t)\dot{h}(t)\}dt.$$

Тем самым доказана формула (17). Непрерывность $\mathcal{J}'(\hat{x}(\cdot))$ следует из непрерывной дифференцируемости оператора Немыцкого (см. Предложение 5.1). □

Лекция 8

В задачах вариационного исчисления и оптимального управления встречаются также интегральные функционалы типа (16) с переменными пределами интегрирования t_0 и t_1. Включим их в общую схему. Пусть предположения относительно $f(t, x, \dot{x})$ те же, что и раньше, $\Delta \subset \mathbb{R}$-некоторый отрезок,

$$\mathcal{U} = \{(x(\cdot), t_0, t_1) : x(\cdot) \in C^1(\Delta, \mathbb{R}^n),\ t_0, t_1 \in \mathrm{int}\Delta;\ (t, x(t), \dot{x}(t)) \in U\ \forall t \in \Delta\}, \tag{20}$$

Определим отображение $\mathcal{J} : W \to \mathbb{R}^m$ равенством

$$\mathcal{J}(x(\cdot), t_0, t_1) = \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t))dt. \tag{21}$$

Предложение 5.4. *Интегральное отображение (21) непрерывно дифференцируемо на множестве (20) и при этом*

$$\mathcal{J}'(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)[h(\cdot), \tau_0, \tau_1] = \int_{t_0}^{t_1} (\hat{f}_x(t)h(t) + \hat{f}_{\dot{x}}(t)\dot{h}(t))dt + \hat{f}(\hat{t}_1)\tau_1 - f(\hat{t}_0)\tau_0, \tag{22}$$

где $\hat{f}_x(t)$ *и* $\hat{f}_{\dot{x}}(t)$ *определены в (18),(19), а*

$$\hat{f}(t) = f(t, \hat{x}(t), \dot{\hat{x}}(t)). \tag{23}$$

Доказательство. Воспользуемся Теоремой 4.4 о полном дифференциале. Частные производные существуют в силу Предложения 5.3 и классической теоремы о производной интеграла по верхнему и нижнему пределам интегрирования:

$$\mathcal{J}_{x(\cdot)}(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)[h(\cdot)] = \int_{t_0}^{t_1} (\hat{f}_x(t)h(t) + \hat{f}_{\dot{x}}(t)\dot{h}(t))dt,$$

$$\mathcal{J}_{t_0}(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)[\tau_0] = -\hat{f}(\hat{t}_0)\tau_0$$

$$\mathcal{J}_{t_1}(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)[\tau_1] = \hat{f}(\hat{t}_1)\tau_1$$

Проверим непрерывность частной производной по t_0:

$$\|\mathcal{J}_{t_0}(x(\cdot), t_0, t_1) - \mathcal{J}_{t_0}(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)\| = \sup_{\|\tau_0\| \le 1} \|f(t_0, x(t_0), \dot{x}(t_0))\tau_0 -$$

$$-f(\hat{t}_0, \hat{x}(t_0), \dot{\hat{x}}(t_0))\tau_0\| \le \|f(t_0, x(t_0), \dot{x}(t_0)) - f(\hat{t}_0, \hat{x}(t_0), \dot{\hat{x}}(t_0))\|. \tag{24}$$

Кроме того,

$$\|x(t_0) - \hat{x}(\hat{t}_0)\| \le \|x(t_0) - \hat{x}(t_0)\| + \|\hat{x}(t_0) - \hat{x}(\hat{t}_0)\| \le \|x(\cdot) - \hat{x}(\cdot)\|_{C^1} + \|\hat{x}(t_0) - \hat{x}(\hat{t}_0)\|$$

$$\|\dot{x}(t_0) - \dot{\hat{x}}(\hat{t}_0)\| \le \|\dot{x}(t_0) - \dot{\hat{x}}(t_0)\| + \|\dot{\hat{x}}(t_0) - \dot{\hat{x}}(\hat{t}_0)\| \le \|x(\cdot) - \hat{x}(\cdot)\|_{C^1} + \|\dot{\hat{x}}(t_0) - \dot{\hat{x}}(\hat{t}_0)\|.$$

Поэтому, если $t_0 \to \hat{t}_0$ и $x(\cdot) \to \hat{x}(\cdot)$ в $C^1(\Lambda, \mathbb{R}^n)$, то $x(t_0) \to x(\hat{t}_0)$, $\dot{x}(t_0) \to \dot{x}(\hat{t}_0)$ и, значит правая часть неравенства (24) стремится к нулю. Следовательно, $\mathcal{J}_{t_0}(x(\cdot), t_0, t_1)$ непрерывно зависит от $(x(\cdot), t_0, t_1)$ (от t_1 эта производная не зависит вовсе). Непрерывность $\mathcal{J}_{t_1}(x(\cdot), t_0, t_1)$ проверяется аналогично, а непрерывность $\mathcal{J}_{x(\cdot)}(x(\cdot), t_0, t_1)$ проверяется как в предложениях 5.3 и 5.1.

Применяя теорему 4.4 о полном дифференциале, получим (22). □

5.4. **Оператор краевых условий.** Пусть функция $\psi(t_0, x_0, t_1, x_1) : V \to \mathbb{R}^s$ непрерывно дифференцируема на открытом множестве $V \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$, и пусть

$$W = \{(x(\cdot), t_0, t_1) : x(\cdot) \in C^1(\Delta, \mathbb{R}^n),\ t_0, t_1 \in \mathrm{int}\Delta, (t_0, x(t_0), t_1, x(t_1)) \in V\} \tag{25}$$

Отображение $\Psi : W \to \mathbb{R}^s$, определяемое равенством

$$\Psi(x(\cdot), t_0, t_1) = \psi(t_0, x(t_0), t_1, x(t_1)), \tag{26}$$

называется оператором краевых условий.

Предложение 5.5. *Оператор краевых условий (26) непрерывно дифференцируем на множестве (25) и при этом*

$$\Psi'(\hat{x}(\cdot), \hat{t}_0, \hat{t}_1)[h(\cdot), \tau_0, \tau_1] = \hat{\psi}_{t_0}\tau_0 + \hat{\psi}_{x_0}[h(\hat{t}_0) + \dot{\hat{x}}(\hat{t}_0)\tau_0] + \hat{\psi}_{t_1}\tau_1 + \hat{\psi}_{x_1}[h(\hat{t}_1) + \dot{\hat{x}}(\hat{t}_1)\tau_1] \tag{27}$$

где

$$\hat{\psi}_{t_j} = \psi_{t_j}(\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1)),\ j = 0, 1 \tag{28}$$

$$\hat{\psi}_{x_j} = \psi_{x_j}(\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1)),\ j = 0, 1 \tag{29}$$

Доказательство. Рассмотрим сначала отображение $\mathrm{ev} : C^1(\Delta, \mathbb{R}^n) \to \mathbb{R}^n$, определяемое равенством $\mathrm{ev}(x(\cdot), t_0) = x(t_0)$. Оно линейно по $x(\cdot)$ и потому
$$\mathrm{ev}_{x(\cdot)}[\hat{x}(\cdot)] = h(\hat{t}_0);$$
частная производная по t_0 - это обычная производная
$$\mathrm{ev}_{t_0}(\hat{x}(\cdot), \hat{t}_0)[\tau_0] = \dot{\hat{x}}(\hat{t}_0)\tau_0.$$
Проверим непрерывность:
$$\|\mathrm{ev}_{x(\cdot)}(x(\cdot), t_0) - \mathrm{ev}_{x(\cdot)}(\hat{x}(\cdot), \hat{t}_0)\| = \sup_{\|h(\cdot)\|_{C^1} \le 1} \|h(t_0) - h(\hat{t}_0)\| \le$$
$$\le \sup_{\|h(\cdot)\|_{C^1}} \sup_{\theta \in [t_0, \hat{t}_0]} \|\dot{h}(\theta)\| |t_0 - \hat{t}_0| \le |t_0 - \hat{t}_0| \to 0$$
при $t_0 \to \hat{t}_0$. Далее
$$\|\mathrm{ev}_{t_0}(x(\cdot), t_0) - \mathrm{ev}_{t_0}(\hat{x}(\cdot), \hat{t}_0)\| = \sup_{|\tau_0| \le 1} \|\dot{x}(t_0)\tau_0 - \dot{\hat{x}}(\hat{t}_0)\tau_0\| \le \|\dot{x}(t_0) - \dot{\hat{x}}(\hat{t}_0)\| \to 0,$$
когда $t_0 \to \hat{t}_0$ и $x(\cdot) \to \hat{x}(\cdot)$ в пространстве $C^1(\Delta, \mathbb{R}^n)$, как это было показано в Предложении 5.4. В силу Теоремы 4.4 о полном дифференциале
$$\mathrm{ev}'(\hat{x}(\cdot), \hat{t}_0)[h(\cdot), \tau_0] = h(\hat{t}_0) + \dot{\hat{x}}(\hat{t})\tau_0$$

С помощью Теоремы 4.1 о суперпозиции проверяется дифференцируемость отображения (26) и справедливость равенства (27). □

6. Теорема Люстерника

Этот параграф посвящен теореме Л.А.Люстерника, имеющей важные приложения в теории экстремальных задач. Начнем с доказательства теоремы, связанной с исследованием свойств правого обратного отображения, построенного в Лемме 3.4 (являющейся обобщением Теоремы 3.3 Банаха об обратном операторе).

Теорема 6.1. *(об оценке расстояния). Пусть X и Y - банаховы пространства U - окрестность точки $\hat{x} \in X$, $F : U \to Y$ - строго дифференцируемое в точке $\hat{x}$ отображение, $ImF'(\hat{x}) = Y$. Тогда существуют константа $K > 0$, окрестности $V \subset U$ и W соответственно точек $\hat{x}$ и $F(\hat{x})$ и отображение $\varphi : V \times W \to X$ такие, что для любых $x \in V$ и $y \in W$ выполнены соотношения $F(x + \varphi(x, y)) = y$ и $\|\varphi(x, y)\|_X \le K\|y - F(x)\|_Y$.*

Доказательство. Обозначим через $B_X(\hat{x}; r)$ - замкнутый, а через $U_X(\hat{x}; r)$ - открытый шар в пространстве X, с центром в точке $\hat{x}$ и радиуса r. Положим $\Lambda = F'(\hat{x})$, и пусть R - правый обратный оператор к Λ. Для $\varepsilon = 1/(2C)$, где C - константа из неравенства в Лемме 3.4, существует такое $\delta > 0$, что $B_X(\hat{x}; \delta) \subset U$, функция F непрерывна на $B_X(\hat{x}; \delta)$ и
$$\|F(x') - F(x'') - \Lambda(x' - x'')\|_Y \le \varepsilon \|x' - x''\|_X \tag{1}$$
для $x', x'' \in U_X(\hat{x}; \delta)$. Положим $W = U_Y(F(\hat{x}); \frac{\delta}{8C})$.

В силу непрерывности отображения F в точке $\hat{x}$ существует такое $\delta_1 > 0$, что $F(x) \in W$, если $x \in U_X(\hat{x}; \delta_1)$. Пусть $\gamma = \min(\delta_1, \frac{\delta}{2})$. Положим $V = U_X(\hat{x}; \gamma)$.

Пусть $x \in V, y \in W$. Рассмотрим последовательность
$$x_{k+1} = x_k + R(y - F(x_k)),\ k = 0, 1, 2, \dots,\ x_0 = x. \tag{2}$$
Докажем по индукции, что $x_k \in U_X(\hat{x}; \delta),\ k = 0, 1, 2, \dots)$

По условию $x_0 \in U_X(\hat{x};\delta)$. Пусть доказано, что элементы x_k при всех $k = 0, 1, \dots, n$ определены и принадлежат $U_X(\hat{x};\delta)$. Согласно формуле (2) имеем

$$\Lambda(x_n - x_{n-1}) = y - F(x_{n-1}). \tag{3}$$

Снова используя формулу (2), оценку для нормы правого обратного и соотношения (3), (1) и затем итерируя процедуру, получим

$$\|x_{n+1} - x_n\|_X \le C\|y - F(x_n)\|_Y = C\|F(x_n) - F(x_{n-1}) - \Lambda(x_n - x_{n-1})\|_Y \le$$
$$\le \frac{1}{2}\|x_n - x_{n-1}\|_X \le \dots \le \frac{1}{2^n}\|x_1 - x\|_X.$$

Отсюда

$$\|x_{n+1} - \hat{x}\|_X = \|x_{n+1} - x_n + x_n - \dots + x_1 - x + x - \hat{x}\|_X \le \left(\frac{1}{2^n} + \dots + \frac{1}{2}\right)\|x_1 - x\|_X +$$
$$+\|x - \hat{x}\|_X \le 2\|x_1 - x\|_X + \|x - \hat{x}\|_X \le 2C\|y - F(x)\|_Y + \|x - \hat{x}\|_X \le 2C\|y - F(\hat{x})\|_Y +$$
$$+2C\|F(\hat{x}) - F(x)\|_Y + \|x - \hat{x}\|_X < 2C\left(\frac{\delta}{8C} + \frac{\delta}{8C}\right) + \frac{\delta}{2} = \delta$$

и тем самым $x_{n+1} \in U_X(\hat{x};\delta)$.

Последовательность $\{x_n\}$ фундаментальна, так как если $m > 0$, то

$$\|x_{n+m} - x_n\|_X \le \|x_{n+m} - x_{n+m-1}\|_X + \dots + \|x_{n+1} - x_n\|_X \le$$
$$\le \left(\frac{1}{2^{m-1}} + \dots + \frac{1}{2} + 1\right)\|x_{n+1} - x_n\|_X \le \frac{2}{2^n}\|x_1 - x\|_X \to 0$$

при $n \to \infty$. Следовательно, в силу полноты пространства X у этой последовательности существует предел. Пусть $\psi(x,y) = \lim_{n\to\infty} x_n$. Поскольку отображение F непрерывно на $B_X(\hat{x};\delta)$, используя равенство (3), получим

$$\|F(\psi(x,y)) - y\|_Y = \lim_{n\to\infty}\|F(x_n) - y\|_Y = 0,$$

т.е. $F(\psi(x,y)) = y$. Положим $\varphi(x,y) = \psi(x,y) - x$. Тогда $F(\varphi(x,y) + x) = y$.

Далее, (как уже неоднократно делалось), воспользуемся неравенствами

$$\|x_n - x\|_X \le \|x_n - x_{n-1}\|_X + \dots + \|x_1 - x\|_X \le 2\|x_1 - x\|_X \le 2C\|y - F(x)\|_Y.$$

Переходя здесь к пределу при $n \to \infty$ и обозначая $K = 2C$, получим, что $\|\varphi(x,y)\| \le K\|y - F(x)\|_Y$. □

Лекция 9

Пусть M - некоторое подмножество нормированного пространства X. Элемент $h \in X$ называется касательным вектором к M в точке $\hat{x}$, если существует такое $\varepsilon > 0$ и такое отображение $r : (-\varepsilon, \varepsilon)$, что

$$\hat{x} + th + r(t) \in M \quad \text{для всех} \quad t \in (-\varepsilon, \varepsilon) \quad \text{и} \quad \frac{\|r(t)\|_X}{|t|} \to 0 \quad \text{при} \quad t \to 0.$$

Множество всех касательных векторов к M в точке $\hat{x}$ обозначим $T_{\hat{x}}M$. Если это множество является подпространством, то его называют касательным пространством (к M в точке $\hat{x}$).

Теорема 6.2. *(Люстерника о касательном пространстве). Пусть X, Y - банаховы пространства, U - окрестность точки $\hat{x} \in X$, отображение $F : U \to Y$ строго дифференцируемо в точке $\hat{x}$, $ImF'(\hat{x}) = Y$ и $M = \{x \in X : F(x) = F(\hat{x})\}$. Тогда $T_{\hat{x}}M = KerF'(\hat{x})$.*

Доказательство. Пусть $h \in T_{\hat{x}}M$. Тогда, используя определение касательного вектора и дифференцируемость отображения F в точке $\hat{x}$, получим

$$F(\hat{x}) = F(\hat{x} + th + r_1(t)) = F(\hat{x}) + tF'(\hat{x})h + r_2(t),$$

где $\frac{\|r_i(t)\|_X}{|t|} \to 0$ при $t \to 0,\ i = 1, 2$. Отсюда следует, что $F'(\hat{x})h = 0$, и, значит, $T_{\hat{x}}M \subset \mathrm{Ker}F'(\hat{x})$.

Обратно, пусть $h \in \mathrm{Ker}F'(\hat{x})$ и φ- функция из Теоремы 6.1 об оценке расстояния. Тогда согласно этой теореме для достаточно малых t имеем

$$F(\hat{x} + th + \varphi(\hat{x} + th, F(\hat{x}))) = F(\hat{x})$$

и

$$\|\varphi(\hat{x} + th, F(\hat{x}))\|_Y \le K\|F(\hat{x}) - F(\hat{x} + th)\|_Y$$

Так как отображение F строго дифференцируемо в точке $\hat{x}$, для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что

$$\|F(\hat{x}) - F(\hat{x} + th)\|_Y = \|F(\hat{x}) - F(\hat{x} + th) + F'(\hat{x})th\|_Y \le \varepsilon|t|\|h\|_X,$$

если $|t| < \delta$. Это означает, что

$$\varphi(\hat{x} + th, F(\hat{x})) = r_3(t), \quad \text{и} \quad \frac{\|r_3(t)\|_X}{|t|} \to 0 \quad \text{при } t \to 0,$$

т.е. $h \in T_{\hat{x}}M$. □

7. Принцип Лагранжа для задач с ограничением типа равенств

7.1. **Постановка задачи.** Сформулируем абстрактную экстремальную задачу с ограничением типа равенств.

Пусть X, Y - банаховы пространства, U - открытое подмножество X, $f : U \to \mathbb{R}$ - функционал, $F : U \to Y$ - отображение. Рассмотрим экстремальную задачу

$$f(x) \to \mathrm{extr}, \tag{1}$$

$$F(x) = 0. \tag{2}$$

Определение 7.1. *Вектор $x \in X$ называется допустимым элементом задачи (1),(2), если $x \in U$ и удовлетворяет (2), т.е. если $x \in M = \{x \in U : F(x) = 0\}$. При этом множество M называется множеством допустимых элементов.*

Определение 7.2. *Допустимый элемент $\hat{x} \in M$ называется точкой локального минимума (максимума) задачи (1),(2), если существует такое $\varepsilon > 0$, что шар $B_X(\hat{x}; \varepsilon) = \{x \in X : \|x - \hat{x}\| < \varepsilon\} \subset U$ и $\forall x \in B_X(\hat{x}; \varepsilon) \cap M$ выполнено неравенство $f(x) \ge f(\hat{x}) \quad (f(x) \le f(\hat{x}))$.*

Предположим, что доказано существование точки $\hat{x}$ локального экстремума задачи (1),(2). Спрашивается, как выписать уравнения, позволяющие определить $\hat{x}$? Этот вопрос обсуждается в следующем пункте.

7.2. **Принцип Лагранжа.** .Чтобы выписать уравнения, позволяющие определить $\hat{x}$, используется принцип Лагранжа, состоящий в следующем: а) составим функцию Лагранжа

$$\mathcal{L}(x, \lambda, y^*) = \lambda f(x) + \langle y^*, F(x)\rangle, \text{ где } \lambda \in \mathbb{R},\ y^* \in Y^*; \tag{3}$$

б) ниже будет показано, что если $\hat{x}$-точка локального экстремума задачи (1),(2), то существует набор множителей Лагранжа $(\lambda, y^*) \in \mathbb{R} \times Y^* \setminus \{0\}$, такой что

$$\mathcal{L}'_x(\hat{x}, \lambda, y^*)[h] = \lambda\langle f'(\hat{x}), h\rangle + \langle y^*, F'(\hat{x})h\rangle = 0, \forall h \in X. \tag{4}$$

Обоснованию этого принципа посвящена следующая теорема:

Теорема 7.1. *(Принцип Лагранжа) Пусть X, Y - банаховы пространства, $U \subset X$ - открытое множество, $\hat{x} \in U$ - решение задачи (1),(2), причем f, F строго дифференцируемы в точке $\hat{x}$, и множество $ImF'(\hat{x})$ замкнуто в Y. Тогда существуют такие множители Лагранжа $(\lambda, y^*) \in \mathbb{R} \times Y^* \setminus \{0\}$, что для функции Лагранжа (3) справедливо соотношение (4). Если $ImF'(\hat{x}) = Y$, то $\lambda \neq 0$, и можно считать, что $\lambda = 1$.*

Доказательство. Рассмотрим оператор $G(x) = (f(x), F(x)) : U \subset X \to \mathbb{R} \times Y$. По Лемме 3.6 о замкнутости образа у его производной, т.е. у оператора $G'(\hat{x})$: $X \to \mathbb{R} \times Y$ образ $\mathrm{Im}G'(\hat{x})$ замкнут в пространстве $\mathbb{R} \times Y$. Действительно, образ оператора $\mathrm{Im}F'(\hat{x})$ замкнут в Y по предположению, а образ оператора $f'(\hat{x}) : \mathrm{Ker}F'(\hat{x}) \to \mathbb{R}$ замкнут в силу линейности оператора и одномерности $\mathbb{R}$.

1) Предположим, что $\mathrm{Im}G'(\hat{x}) \subsetneqq \mathbb{R} \times Y$. По Лемме 3.5 о нетривиальности аннулятора существует ненулевой функционал $\Lambda \in (\mathbb{R} \times Y)^*$ такой, что $\forall h \in X$ $\langle \Lambda, G'(\hat{x})h\rangle = 0$. По лемме 3.1 о виде линейного функционала на прямом произведении пространств, существуют такие $\lambda \in \mathbb{R}, y^* \in Y^*$ что

$$0 = \langle \Lambda, G'(\hat{x})h\rangle = \lambda f'(\hat{x}) + \langle y^*, F'(\hat{x})h\rangle = \mathcal{L}'_x(\hat{x}, \lambda, y^*)[h],$$

т.е. справедливо равенство (4). Пусть $\mathrm{Im}F'(\hat{x}) = Y$. Если допустить, что $\lambda = 0$, то $\langle y^*, F'(\hat{x})h\rangle = 0\ \forall h \in X$, что невозможно. Значит $\lambda \neq 0$, и можно считать, что $\lambda = 1$

2) Нам осталось доказать, что соотношение

$$\mathrm{Im}G'(\hat{x}) = \mathbb{R} \times Y \tag{5}$$

невозможно, если производная берется в точке локального экстремума $\hat{x}$. Допустим противное, т.е. что соотношение (5) справедливо. Тогда $\mathrm{Im}F'(\hat{x}) = Y$, т.е. для любого $\hat{y} \in Y$ существует такой вектор $\hat{h} \in X$, что $F'(\hat{x})[\hat{h}] = \hat{y}$. Обозначим через $\hat{r}$ вещественное число, определяемое равенством $\langle f'(\hat{x}), \hat{h}\rangle = \hat{r}$. Для доказательства (5) нужно было бы для любого $\tilde{r} \in \mathbb{R}$ найти такое $\tilde{h} \in X$, что

$$\langle f'(\hat{x}), \tilde{h}\rangle = \tilde{r},\ F'(\hat{x})[\tilde{h}] = \hat{y}. \tag{6}$$

Однако мы покажем, что из второго из равенств (6) следует, что $\langle f'(\hat{x}), \tilde{h}\rangle = \hat{r}$. Отметим, что $\tilde{h} = \hat{h} + h$, где $h \in \mathrm{Ker}F'(\hat{x})$. Пусть $M = \{x \in U : F(x) = F(\hat{x})\}$. По теореме 6.1 Люстерника о касательном пространстве $T_{\hat{x}}M = \mathrm{Ker}F'(\hat{x})$ и значит для любого $h \in \mathrm{Ker}F'(\hat{x})$ существует такое отображение $\alpha(\lambda)$, что $\|\alpha(\lambda)\|/|\lambda| \to 0$ при $\lambda \to 0$ и значение $\lambda = 0$ является точкой экстремума задачи

$$\varphi(\lambda) = f(\hat{x} + \lambda h + \alpha(\lambda)) \to \mathrm{extr}, \quad F(\hat{x} + \lambda h + \alpha(\lambda)) = 0. \tag{7}$$

Из (7) следует, что $\varphi'(0) = 0 = \langle f'(\hat{x}), h\rangle$, причем, напомним, что второе из этих равенств справедливо для любого $h \in \mathrm{Ker}F'(\hat{x})$. Это доказывает, что в соотношении (6) $\tilde{r}$ однозначно определяется по $\hat{y}$. Отсюда следует невозможность соотношения (5). □

8. Принцип Лагранжа для выпуклых задач

8.1. **Выпуклые функции и множества.** Напомним простейшие определения.

Определение 8.1. *а)Пусть X -линейное пространство. Подмножество $U \subset X$ называется выпуклым, если для любых $x_1, x_2 \in U$ и любого $\alpha \in [0,1]$ $\alpha x_1 + (1-\alpha)x_2 \in U$*

б)Пусть функция $f(x)$ определена на выпуклом множестве U. Функция f называется выпуклой, если для любых $x_1, x_2 \in U$ и $\alpha \in [0,1]$ она удовлетворяет неравенству $f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha)f(x_2)$ (неравенству Иенссена).

В теории выпуклых функций часто рассматривают функции

$$f : X \to \overline{\mathbb{R}} := \{-\infty\} \cup \mathbb{R} \cup \{\infty\}.$$

Для таких функций важным оказывается множество

$$\mathrm{dom}f := \{x \in X : f(x) < \infty\},$$

которое называется эффективным множеством. Часто будет также использоваться множество

$$\mathrm{epi}f = \{(\alpha, x) \in \mathbb{R} \times \mathrm{dom}f : \alpha \ge f(x)\},$$

которое называется надграфиком f.

Определение 8.2. *а)Функция $f : X \to \overline{\mathbb{R}}$ называется выпуклой, если множество $epif$ является выпуклым в $\mathbb{R} \times domf$.*

б)Функция $f : X \to \overline{\mathbb{R}}$ называется собственной, если $domf \ne \emptyset$ и $f(x) > -\infty\ \forall x \in X$.

В дальнейшем мы будем рассматривать только собственные функции.

Предложение 8.1. *Пусть f - собственная функция. Тогда функция $f : X \to \overline{\mathbb{R}}$ выпукла в том и только том случае, если*

$$\forall x_1, x_2 \in domf,\ \forall \lambda \in [0,1] \quad f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda)f(x_2). \quad (1)$$

Доказательство. Пусть функция f выпукла (в смысле Определения 8.2a)), и значит для любых $(\alpha_1, x_1), (\alpha_2, x_2) \in \mathrm{epi}f$ (т.е. $\forall x_1, x_2 \in \mathrm{dom}f$, $\forall \alpha_1 \ge f(x_1), \alpha_2 \ge f(x_2)$) при любом $\lambda \in [0,1]$ справедливо включение $\lambda(\alpha_1, x_1) + (1-\lambda)(\alpha_2, x_2) \in \mathrm{epi}f$, , то есть выполнено неравенство

$$\lambda\alpha_1 + (1-\lambda)\alpha_2 \ge f(\lambda x_1 + (1-\lambda)x_2)$$

Взяв в последнем неравенстве $\alpha_1 = f(x_1), \alpha_2 = f(x_2)$, получим справедливость неравенства (1).

Предположим теперь, что выполнено условие (1). Из неравенства в этом условии сразу следует выпуклость множества $\mathrm{dom}f$. Пусть теперь $\lambda \in [0,1], \infty > \alpha_1 \ge f(x_1), \infty > \alpha_2 > f(x_2)$ т.е. $(\alpha_1, x_1), (\alpha_2, x_2) \in \mathrm{epi}f$. Из этих неравенств и (1) следует что $\lambda\alpha_1 + (1-\lambda)\alpha_2 \ge \lambda f(x_1) + (1-\lambda)f(x_2) \ge f(\lambda x_1 + (1-\lambda)x_2)$ и значит $(\lambda\alpha_1 + (1-\lambda)\alpha_2, \lambda x_1 + (1-\lambda)x_2) \in \mathrm{epi} f$. Тем самым установлена выпуклость множества $\mathrm{epi} f$, а значит и выпуклость функции f в смысле Определения 8.2а). $\square$

Лекция 10

8.2. **Конечномерная теорема отделимости.** Прежде, чем переходить к изложению теории выпуклых экстремальных задач, сформулируем теорему, которая будет использована в этой теории.

Теорема 8.1. *(Конечномерная теорема отделимости) Пусть $\mathbb{R}^m \supset C$ - выпуклое множество, и $0 \notin C$. Тогда существует такой вектор $(\lambda_1, \ldots, \lambda_m) \neq 0$, что $\sum_{j=1}^m \lambda_j c_j \geq 0$ для любого вектора $c = (c_1, \ldots, c_m) \in C$.*

Доказательство этой теоремы можно найти в параграфе 1.3 главы 1 книги В.М.Алексеева, В.М.Тихомирова, С.В.Фомина "Оптимальное управление".

8.3. **Выпуклая экстремальная задача.** Пусть X-линейное пространство, $f_j : X \to \overline{\mathbb{R}},\ j = 0, 1, \ldots, m$- собственные выпуклые функции, $X \supset A$ - выпуклое множество. Рассмотрим задачу, которая называется выпуклой экстремальной задачей:

$$f_0(x) \to \min, \tag{2}$$

$$f_j(x) \leq 0,\ j = 1, \ldots, m, \tag{3}$$

$$x \in A, \tag{4}$$

Определение 8.3. *Вектор $x \in X$ называется допустимым элементом задачи (2)-(4), если x удовлетворяет соотношениям (3),(4). Символом $\mathcal{A}$ будет обозначаться множество всех допустимых элементов этой задачи.*

Определение 8.4. *Вектор $\hat{x}$ называется решением задачи (2)-(4), если для любого $x \in \mathcal{A}$ $f_0(x) \geq f_0(\hat{x})$.*

Отметим, что для выпуклых экстремальных задач естественным является определение именно минимума, но не максимума. Покажем, что в теории выпуклых экстремальных задач фактически отсутствует понятие локального минимума.

Предложение 8.2. *Пусть линейное пространство X является нормированным (с нормой $\|\cdot\|$), а $\hat{x}$ - точка локального минимума задачи (2)-(4) (т.е. $\exists \varepsilon > 0\ \forall x \in \mathcal{A} : \|x - \hat{x}\| < \varepsilon$ справедливо неравенство $f_0(x) \geq f_0(\hat{x})$). Тогда $\hat{x}$ является точкой абсолютного минимума задачи (2)-(4).*

Доказательство. Нам нужно доказать, что для любого допустимого x справедливо неравенство $f_0(x) \geq f_0(\hat{x})$. Покажем, что при достаточно малом $\alpha \in [0, 1]$ вектор $\tilde{x} = \alpha x + (1-\alpha)\hat{x}$ принадлежит шару $B_X(\hat{x}; \varepsilon)$ в пространстве X с центром в $\hat{x}$ радиуса ε. Действительно, $\|\tilde{x} - \hat{x}\| = \|\alpha x + (1-\alpha)\hat{x} - \hat{x}\| =$

$\alpha\|x-\hat{x}\| < \varepsilon$. Так как $\hat{x}$ - точка локального минимума, то $f_0(\hat{x}) \le f_0(\tilde{x}) = f_0(\alpha x + (1-\alpha)\hat{x}) \le \alpha f(x) + (1-\alpha) f(\hat{x})$, откуда $f_0(\hat{x}) \le f_0(x)$. □

8.4. **Принцип Лагранжа для выпуклой задачи.** Функция Лагранжа для задачи (2)-(4) имеет вид:

$$\mathcal{L}(x,\lambda) = \mathcal{L}(x, \lambda_0, \ldots, \lambda_m) = \sum_{j=0}^{m} \lambda_j f_j(x)$$

Принцип Лагранжа для указанной задачи формулируется следующим образом: если $\hat{x}$ - точка минимума задачи (2)-(4), то существует такой ненулевой вектор $\lambda = (\lambda_0, \ldots, \lambda_m)$, что $\hat{x}$ является также точкой минимума задачи

$$\mathcal{L}(x,\lambda) \to \min, \quad x \in A.$$

Другими словами принцип Лагранжа является принципом снятия ограничений $f_j(x) \le 0, \ j = 1, \ldots m$, т.е. он сводит решение задачи (2)-(4) к решению аналогичной задачи, но без ограничений (3).

Обоснование принципа Лагранжа дается в следующей теореме:

Теорема 8.2. *(Кун-Таккер) 1) Пусть X - линейное пространство, $f_j(x)$, $j = 0, \ldots, m$ - собственные выпуклые функции, $X \supset A$ - выпуклое множество, $\hat{x}$ - точка минимума задачи (2)-(4). Тогда существует вектор $\lambda = (\lambda_0, \ldots, \lambda_m) \neq 0$, удовлетворяющий следующим условиям:*

$$a)\lambda_j \ge 0, \ j = 0, \ldots, m \quad \textit{(условие неотрицательности)},$$
$$b)\lambda_j f_j(\hat{x}) = 0, \ j = 1, \ldots, m \quad \textit{(условие дополняющей нежёсткости)},$$
$$c)\mathcal{L}(\hat{x},\lambda) \le \mathcal{L}(x,\lambda) \ \forall x \in \mathcal{A} \quad \textit{(условие минимума)}.$$

2)Если $\lambda_0 \neq 0$, то условия a),b),c) достаточны для того, чтобы допустимая точка $\hat{x}$ была решением задачи (2)-(4).

3) Для выполнения условия $\lambda_0 \neq 0$ достаточно, чтобы нашлась точка $\tilde{x} \in \mathcal{A}$, в которой выполнено условие Слейтера, а именно: $f_j(\tilde{x}) < 0, \ j = 1, \ldots, m$.

Доказательство. Будем считать, что для решения $\hat{x}$ в задаче (2)-(4) выполнено равенство $f_0(\hat{x}) = 0$. Этого легко добиться, заменив в задаче (2)-(4) функцию $f_0(x)$ на функцию $x \to f_0(x) - f_0(\hat{x})$.

Введем в пространстве $\mathbb{R}^m$ множество

$$C = \{\alpha = (\alpha_0, \ldots, \alpha_m) : \exists x \in \mathcal{A} \ \ \alpha_0 > f_0(x), \ \alpha_j \ge f_j(\hat{x}), \ j = 1, \ldots, m\}.$$

Покажем, что множество C выпукло. Действительно, пусть $\alpha, \beta \in C$ и значит $\exists x_1 \in \mathcal{A} : \alpha_0 > f_0(x_1), \ \ \alpha_j \ge f_j(x_1)$ и $\exists x_2 \in \mathcal{A} : \beta_0 > f_0(x_2), \ \ \beta_j \ge f_j(x_2)$ для $j = 1, \ldots, m$. Тогда при любом $\lambda \in [0,1]$ справедливо неравенство $f_0(\lambda x_1 + (1-\lambda)x_2) \le \lambda f_0(x_1) + (1-\lambda) f_0(x_2) < \lambda\alpha_0 + (1-\lambda)\beta_0$, а также аналогичные неравенства для f_j с заменой знака $<$ на $\le$. Это доказывает выпуклость множества C.

Покажем, что $0 \notin C$. Действительно, если $0 = (0, \ldots, 0) \in C$, то существует такой $x \in \mathcal{A}$, что $0 > f_0(x), \ 0 \ge f_j(x)$, и значит $f_0(x) < f_0(\hat{x}) = 0$, откуда следует, что $\hat{x}$ не является точкой минимума задачи (2)-(4).

Применяя к множеству C конечномерную теорему отделимости (см. Теорему 8.1) получим, что

$$\exists \lambda = (\lambda_0, \ldots, \lambda_m) \neq 0, \text{ что } (\lambda, \alpha) = \sum_{j=0}^{m} \lambda_j \alpha_j \ge 0 \ \forall \alpha = (\alpha_0, \ldots, \alpha_m) \in C \qquad (5)$$

1) Покажем, что набор $\lambda = (\lambda_j,\ j = 0, \dots, m)$ из (5) удовлетворяет условиям a),b),c) из формулировки теоремы.

a) (Условие неотрицательности.) Очевидно, $\hat{\alpha} = (1, 0, \dots, 0) \in C$, так как $1 > f_0(\hat{x}) = 0, 0 \geq f_j(\hat{x}), j = 1, \dots, m$ (силу соотношения (2)). Поэтому $\lambda_0 = (\lambda, \hat{\alpha}) \geq 0$. Точно также $\tilde{\alpha} = (\delta, 0, \dots, 1, 0, \dots) \in C$ где $\delta > 0$, а на остальных местах, везде стоят нули, кроме одного места с номером j, где стоит 1. Поэтому $(\lambda, \tilde{\alpha}) = \lambda_0 \delta + \lambda_j \geq 0$, и значит $\lambda_j \geq -\delta\lambda_0 \to 0$ при $\delta \to 0$, т.е. $\lambda_j \geq 0$.

b) (Условие дополняющей нежёсткости.) Очевидно $\lambda_0 f_0(\hat{x}) = 0$, так как $f_0(\hat{x}) = 0$. Если $f_j(\hat{x}) = 0$, то и $\lambda_j f_j(\hat{x}) = 0$.

Пусть $f_j(\hat{x}) < 0$, тогда $(\delta, 0, \dots, 0, f_j(\hat{x}), 0, \dots, 0) \in C$ (здесь $\delta > 0$, и f_j стоит на j-ом месте), и значит $\lambda_0 \delta + \lambda_j f_j(\hat{x}) \geq 0$, откуда $\lambda_j f_j(\hat{x}) \geq -\lambda_0 \delta \to 0$ при $\delta \to 0$. Следовательно, $\lambda_j f_j(\hat{x}) \geq 0$. Учитывая, что $f_j(\hat{x}) < 0$ и в силу a) $\lambda_j \geq 0$, получаем, что $\lambda_j = 0$ и $\lambda_j f_j(\hat{x}) = 0$.

c) (Условие минимума.) Очевидно, $\alpha = (\delta + f_0(x), f_j(x), j = 1, \dots, m) \in C$, где $\delta > 0, x \in \mathcal{A}$, и поэтому $\lambda_0 \delta + \sum_{j=0}^m \lambda_j f_j(x) \geq 0$. Следовательно, $\sum_{j=0}^m \lambda_j f_j(x) \geq (-\delta\lambda_0) \to 0)$ при $\delta \to 0$. Это доказывает требуемое неравенство, так как в силу условия дополняющей нежёсткости b) $0 = \sum_{j=0}^m \lambda_j f_j(\hat{x})$.

2)Докажем при $\lambda_0 > 0$ достаточность условий a),b),c), чтобы допустимый элемент $\hat{x}$ был решением задачи (2)-(4). Пусть $\lambda_0 > 0$, $\hat{x}$ - допустимый элемент задачи (2)-(4), для которого выполнены условия a),b),c) и x - произвольный допустимый элемент. Тогда справедливы соотношения

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

причем первое равенство справедливо в силу условия b), следующее неравенство в силу c), а последнее неравенство - в силу a). Следовательно $\hat{x}$ -точка минимума задачи (2)-(4).

3) Докажем, что из условия Слейтера следует, что $\lambda_0 > 0$. Пусть $\lambda_0 \geq 0$ и существует такой допустимый элемент x задачи (2)-(4), что $f_j(x) < 0,\ j = 1, \dots, m$. Надо доказать , что $\lambda_0 > 0$. Допустим противное, т.е. что $\lambda_0 = 0$. Тогда справедливы соотношения

$$\mathcal{L}(\hat{x}, \lambda) = \sum_{j=1}^{m} \lambda_j f_j(\hat{x}) = 0 > \sum_{j=1}^{m} = \mathcal{L}(x, \lambda),$$

причем первое равенство следует из $\lambda_0 = 0$, второе - из условия дополняющей нежёсткости b), третье соотношение из условия Слейтера. В итоге мы получили неравенство, противоречащее условию минимума c). Следовательно $\lambda_0 > 0$. □

9. Примеры применения принципа Лагранжа

9.1. **Конечномерная задача.** Мы начнем с примеров применения принципа Лагранжа для гладких задач, которым принадлежит и конечномерная задача. Пусть $f_j(x) \in C^1(\mathbb{R}^n)$. Рассмотрим задачу

$$f_0(x) \to \text{extr}, \tag{1}$$

$$f_j(x) = 0,\ j = 1, \dots, m \tag{2}$$

Определение 9.1. *Вектор $x \in \mathbb{R}^n$ называется допустимым в задаче (1),(2), если он удовлетворяет соотношениям (2). Через $\mathcal{A}$ обозначается множество всех допустимых векторов.*

Определение 9.2. *Вектор $\hat{x} \in \mathcal{A}$ называется точкой локального минимума (максимума), если*

$$\exists \varepsilon > 0\, \forall x \in \mathcal{A}: \ \|\hat{x} - x\| < \varepsilon \quad f_0(\hat{x}) \le f_0(x) \quad (f_0(\hat{x}) \ge f_0(x))$$

Точка локального экстремума - это точка либо локального минимума, либо локального максимума.

Функцией Лагранжа задачи (1),(2) называется функция

$$\mathcal{L}(x, \lambda) = \sum_{j=0}^{m} \lambda_j f_j(x), \quad \text{где} \quad \lambda = (\lambda_0, \ldots, \lambda_m).$$

.

Теорема 9.1. *Пусть $f_j(x) \in C^1(\mathbb{R}^n)$, $j = 0, \ldots, m$, а $\hat{x}$ -точка локального экстремума в задаче (1),(2). Тогда*

1) Существует такой вектор $\lambda = (\lambda_0, \ldots, \lambda_m) \neq 0$, что

$$\partial_{x_i} \mathcal{L}(\hat{x}, \lambda) = \sum_{j=0}^{m} \lambda_j \partial_{x_i} f_j(x) = 0, \ i = 1, \ldots, n. \tag{3}$$

2)Пусть $F(x) = (f_1(x), \ldots, f_m(x)), x = (x_1, \ldots, x_n), \nabla = (\partial_{x_1}, \ldots, \partial_{x_n})$. Если отображение $\nabla F(\hat{x}) : \mathbb{R}^n \to \mathbb{R}^m$ является отображением на, то координата λ_0 вектора $\lambda = (\lambda_0, \ldots, \lambda_m)$ отлична от 0.

Доказательство. Доказательство сводится к применению Теоремы 7.1 о принципе Лагранжа для гладких задач с ограничениями типа равенств к нашему случаю. Мы должны взять $X = \mathbb{R}^n$, $Y = \mathbb{R}^m, f(x) = f_0(x)$, $F(x) = (f_1(x), \ldots, f_m(x))$, предположить существование локального экстремума $\hat{x}$. Тогда предположение Теоремы 7.1 о строгой дифференцируемости f, F в точке $\hat{x}$ следует из включения $f_j(x) \in C^1(\mathbb{R}^n)$, $j = 1, \ldots, m$, а предположение о замкнутости $\mathrm{Im} \nabla F(\hat{x})$ в Y следует из конечномерности этих пространств. Поэтому доказываемая теорема следует из теоремы 7.1. □

Теорема 9.1 - это один из простейших частных случаев применения принципа Лагранжа для гладких задач с ограничениями типа равенств. На нем предельно ясно видна важность пункта 2) этой теоремы. Действительно, результатом теоремы является система уравнений, которой удовлетворяет точка локального минимума задачи (1),(2). Она состоит из уравнений (2),(3), то есть из $m+n$ уравнений, которые имеют $n+m+1$ неизвестных $x_1, \ldots, x_n, \lambda_0, \ldots, \lambda_m$. Утверждение 2) Теоремы 9.1 позволяет, поделив (3) на λ_0 и переобозначив $\lambda_j/\lambda_0 = \mu_j$, получить систему уравнений, у которой число уравнений и неизвестных совпадает. Это делает возможным (при соответствующих условиях) однозначное определение локального экстремума из системы (2)(3).

Лекция 11

9.2. **Задача Лагранжа.** Так называется следующая экстремальная задача:

$$J(x, u) = \int_{t_0}^{t_1} f(t, x(t), u(t)) dt \to \mathrm{extr}, \tag{4}$$

$$\dot{x}(t) + \varphi(t, x(t), u(t)) = 0, \tag{5}$$

$$g(x(t_0), x(t_1)) = 0, \tag{6}$$

где

$$f \in C^1([t_0,t_1]\times\mathbb{R}^{n+m}), \quad \varphi \in C^1([t_0,t_1]\times\mathbb{R}^{n+m};\mathbb{R}^n), \quad g \in C^1(\mathbb{R}^{2n};\mathbb{R}^k), \quad (7)$$

-заданные функции, а искомыми функциями являются фазовая функция

$$x(t) \in C^1([t_0,t_1];\mathbb{R}^n), \qquad \text{и управление} \quad u(t) \in C([t_0,t_1];\mathbb{R}^m). \qquad (8)$$

(Физический смысл: обычно $x(t)$ -это координаты нескольких точек, движущихся посредством силы $u(t)$ на основе закона Ньютона (5). Требуется выбрать такое управление, чтобы при выполнении (5) и краевых условий (6) минимизировать (максимизировать) функционал (4).

Определение 9.3. *Пара $(x(t),u(t))$ из (8) называется допустимой, если она удовлетворяет соотношениям (5),(6). Через $\mathcal{A}$ обозначается множество всех допустимых пар.*

Определение 9.4. *Пара $(\hat{x}(t),\hat{u}(t)) \in \mathcal{A}$ называется локальным минимумом (максимумом) задачи (4)-(6), если существует такое $\varepsilon > 0$, что*

$$\forall (x,u) \in \mathcal{A} : \|\hat{x}-x\|_{C^1} + \|\hat{u}-u\|_C < \varepsilon \qquad J(\hat{x},\hat{u}) \le J(x,u) \ (J(\hat{x},\hat{u}) \ge J(x,u)).$$

Локальным экстремумом называется локальный минимум либо локальный максимум.

Применим (формально) принцип Лагранжа к задаче (4)-(6). Соответствующая функция Лагранжа имеет вид

$$\mathcal{L}(x,u,\lambda,p,\nu) = \int_{t_0}^{t_1} L(t,x,\dot{x},u,\lambda,p)dt + l(x(t_0),x(t_1),\nu),$$

где

$$L(t,x,\dot{x},u,\lambda,p) = \lambda f(t,x(t),u(t)) + \langle p(t), \dot{x}(t) + \varphi(t,x(t),u(t))\rangle,$$

$$l(x(t_0),x(t_1),\nu) = \langle \nu, g(x(t_0),x(t_1))\rangle$$

Согласно принципу Лагранжа, если $(\hat{x},\hat{u})$ является точкой локального экстремума задачи (4)-(6), то при некотором наборе $(\lambda,p(\cdot),\nu) \neq 0$ эта же пара является точкой локального экстремума для задачи $\mathcal{L}(t,x,u,\lambda,p,\nu) \to \text{extr}$. Отсюда, очевидно следует, что тогда $\hat{x}$ является точкой локального экстремума для задачи $\mathcal{L}(t,x,\hat{u},\lambda,p,\nu) \to \text{extr}$, а $\hat{u}$ -точкой локального экстремума для задачи $\mathcal{L}(t,\hat{x},u,\lambda,p,\nu) \to \text{extr}$. Отметим, что первая из указанных экстремальных задач (при фиксированном $\hat{u}$) является задачей Больца по x и для неё легко выписать уравнение Эйлера и условие трансверсальности:

$$-\dot{p}(t) + (\varphi_x(t,\hat{x}(t),\hat{u}(t)))^* p(t) = -\lambda f_x(t,\hat{x}(t),\hat{u}(t)), \qquad (9)$$

$$\dot{p}(t_i) = (-1)^i \langle \nu, g_{x(t_i)}(\hat{x}(t_0),\hat{x}(t_1))\rangle, \ i = 0,1 \qquad (10)$$

Необходимым условием для второй из указанных экстремальных задач (при фиксированном $\hat{x}$), очевидно, является уравнение $\hat{\mathcal{L}}_u = 0$, или в более подробной записи

$$\lambda f_u(t,\hat{x}(t),\hat{u}(t)) + (\varphi_u(t,\hat{x}(t),\hat{u}(t))^* p(t) = 0. \qquad (11)$$

Теперь мы готовы сформулировать теорему о принципе Лагранжа для задачи Лагранжа (4)-(6).

Теорема 9.2. *Пусть f, φ, g - функции, определенные соотношениями (7), $(\hat{x}(t), \hat{u}(t))$ -решение задачи (4)-(6). Тогда существуют*

$$(\lambda, p, \nu) \in \{\mathbb{R} \times C^1([t_0, t_1]; \mathbb{R}^n) \times \mathbb{R}^k\} \setminus \{0\}$$

такие, что выполнены соотношения (9)-(11).

Доказательство. Напомним формулировку Теоремы 7.1 о принципе Лагранжа для абстрактной экстремальной задачи с ограничениями типа равенств $j(z) \to \mathrm{extr}$, $F(z) = 0$, где $j : X \to \mathbb{R}$, $F : X \to Y$, (X, Y) - банаховы пространства. Если $\hat{z}$ - решение указанной задачи, j, F строго дифференцируемы в точке $\hat{z}$, а образ отображения $F'(\hat{z}) : X \to Y$ замкнут, то существует такой ненулевой набор $(\lambda, y^*) \in \mathbb{R} \times Y^*$, что $\hat{z}$ является решением задачи $\mathcal{L}'_z(\hat{z}, \lambda, y^*)[h] = 0 \; \forall h \in X$, где $\mathcal{L}(z, \lambda, y^*) = \lambda j(z) + \langle y^*, F(z) \rangle$.

Применяя теорему 7.1 к случаю задачи Лагранжа, получим, что

$$(x(t), u(t)) := z(t) \in C^1([t_0, t_1]; \mathbb{R}^n) \times C([t_0, t_1]; \mathbb{R}^m) := X,$$

функционал $j(z) = J(x, u) : X \to \mathbb{R}$ - непрерывно дифференцируем (что доказывается с помощью оператора Немыцкого), а значит и строго дифференцируем в точке $\hat{z}$; отображение

$$F(z) = (\dot{x}(t) + \varphi(t, x(t), u(t)), g(x(t_0), x(t_1))) : X \to Y := C([t_0, t_1]; \mathbb{R}^n) \times \mathbb{R}^k \quad (12)$$

по тем же причинам строго дифференцируемо в $\hat{z}$. Проверим замкнутость образа оператора $F'(\hat{z}) : X \to Y$. Отметим, что

$$F'(\hat{z})[h, v] = (\dot{h}(t) + \hat{\varphi}'_x(t)h(t) + \hat{\varphi}'_u(t)v(t), \hat{g}'_{x(t_0)}h(t_0) + \hat{g}'_{x(t_1)}h(t_1)),$$

причем образ первой компоненты в правой части этого равенства замкнут и даже действует на все пространство $C([t_0, t_1]; \mathbb{R}^n)$, так как при любом $f(t) \in C([t_0, t_1]; \mathbb{R}^n)$ существует решение задачи $\dot{h} + \hat{\varphi}_x(t)h(t) + \hat{\varphi}_u(t)v(t) = f(t)$, $h|_{t=t_0} = 0$. Обозначим эту компоненту оператора $F'(\hat{z})$ через A. Вторая компонента правой части этого равенства имеет конечномерный образ и поэтому, обозначив этот оператор через B, получим замкнутость пространства $B\mathrm{Ker}A$. Следовательно в силу леммы 3.6 о замкнутости образа, получим, что образ оператора $F'(\hat{z})$ замкнут в Y.

Таким образом все предположения теоремы 7.1 выполнены в нашем случае, а значит верно и утверждение этой теоремы. Чтобы его сформулировать, отметим, что согласно (12), сопряженным к Y является пространство $Y^* = C([t_0, t_1]; \mathbb{R}^n)^* \times \mathbb{R}^k$. Значит ненулевой набор функционалов, задающих функцию Лагранжа для задачи Лагранжа имеет вид $(\lambda, y^*) = (\lambda, \mu^*, \nu) \in \mathbb{R} \times C([t_0, t_1]; \mathbb{R}^n)^* \times \mathbb{R}^k$. Структура функционалов λ, ν очевидна. Обсудим структуру функционала μ^*.

Рассмотрим сначала случай скалярных функций (т.е. функций со значениями в $\mathbb{R}$). Напомним некоторые факты. [1]

[1]См. подробнее в 1)А.Н.Колмогоров,С.В.Фомин, Элементы теории функций и функционального анализа. (гл. 2,6),2) В.М.Алексеев,В.М.Тихомиров, С.В.Фомин, Оптимальное управление (гл. 2.1.9)

Определение 9.5. *Функция $f(t),\ t \in [a,b]$ называется функцией с ограниченной вариацией, если существует такая константа C, что при любом разбиении отрезка $[a,b]$ точками $a = t_0 < t_1, \dots < t_n = b$ выполнено неравенство*

$$\sum_{k=1}^{n} |f(t_k) - f(t_{k-1})| \le C. \tag{13}$$

Определение 9.6. *Функция ограниченной вариации $v(t),\ t \in [a,b]$ называется канонической, если она непрерывна справа во всех точках интервала (a,b) и $v(a) = 0$.*

Теорема 9.3. *(Ф.Рисс) Каждому непрерывному линейному функционалу $x^* \in C([a,b])^*$ соответствует каноническая функция ограниченной вариации $v(\cdot) : [a,b] \to \mathbb{R}$ такая, что для всех $x(\cdot) \in C([a,b])$*

$$\langle x^*, x(\cdot)\rangle = \int_a^b x(t)dv(t),$$

и такое представление единственно: если для канонической $v(\cdot)$

$$\int_a^b x(t)dv(t) = 0,\ \forall x(\cdot) \in C([a,b]), \quad \text{то} \quad v(t) \equiv 0$$

(Отметим, что интегралы в формулировке теоремы Рисса являются интегралами Лебега-Стилтьеса.) Эта теорема легко обобщается на векторный случай. Пусть $e_1 = (1,0,\dots,0), e_2 = (0,1,0,\dots,0),\dots$ - единичные векторы стандартного базиса в $\mathbb{R}^n$. Тогда $x(\cdot) = (x_1(\cdot),\dots,x_n(\cdot)) \in C([a,b];\mathbb{R}^n)$ представляется в виде

$$x(\cdot) = \sum_{k=1}^{n} x_k(\cdot)e_k.$$

Если $\mu^* \in C([a,b];\mathbb{R}^n)^*$, то

$$\langle \mu^*, x(\cdot)\rangle = \sum_{k=1}^{n}\langle \mu^*, x_k(\cdot)e_k\rangle = \sum_{k=1}^{n}\langle x_k^*, x_k(\cdot)\rangle$$

где функционалы $x_k^* \in C([a,b])^*$ определяются равенствами $\langle x_k^*, \xi(\cdot)\rangle = \langle \mu^*, \xi(\cdot)e_k\rangle$. Применяя к x_k^* теорему Рисса, получим представление

$$\langle \mu^*, x(\cdot)\rangle = \sum_{k=1}^{n}\int_a^b x_k(t)dv_k(t). \tag{14}$$

Набор функций ограниченной вариации $\mu(\cdot) = (v_1(\cdot),\dots,v_n(\cdot))$ естественно назвать векторнозначной функцией ограниченной вариации $\mu(\cdot) : [a,b] \to \mathbb{R}^{n*}$. Тогда формула (14) может быть переписана в виде

$$\langle \mu^*, x(\cdot)\rangle = \int_a^b \langle d\mu(t), x(t)\rangle. \tag{15}$$

Как и в теореме Рисса, соответствие между μ^* и $\mu(\cdot)$ будет взаимно однозначным при соблюдении условия "каноничности"функции $\mu(\cdot)$.

Таким образом установлена искомая структура функционала μ^*, и мы можем выписать функцию Лагранжа для задачи Лагранжа:

$$\mathcal{L}(x,u,\lambda,\mu^*,\nu) = \int_{t_0}^{t_1} \lambda f(t,x(t),u(t))dt +$$

$$+\int_{t_0}^{t_1}\langle d\mu(t),(\dot{x}(t)+\varphi(t,x(t),u(t)))\rangle+\langle\nu,g(x(t_0),x(t_1))\rangle. \qquad (16)$$

Мы доказали с помощью теоремы 7.1, что при некотором ненулевом наборе $(\lambda,\mu(\cdot),\nu)$ выполнены соотношения $\mathcal{L}'_x=0,\mathcal{L}'_u=0$. Выпишем первое из них:

$$\langle\hat{\mathcal{L}}'_x,h\rangle=\int_{t_0}^{t_1}\langle\hat{f}'_x(t),h(t)\rangle dt+$$

$$+\int_{t_0}^{t_1}\langle d\mu(t),(\dot{h}(t)+\hat{\varphi}'_x(t)h(t))\rangle+\langle\nu,(\hat{g}'_{x(t_0)}h(t_0)+\hat{g}'_{x(t_1)}h(t_1))\rangle=0. \qquad (17)$$

Лекция 12

Чтобы вывести необходимые условия (9)-(11), нам в первую очередь нужно доказать достаточную гладкость функции ограниченной вариации $\mu(t)$ из (17). Для этого преобразуем (17) с помощью интегрирования по частям:

$$\int_{t_0}^{t_1}\lambda\langle\hat{f}'_x(t),h(t)\rangle dt=-\lambda\langle h(t),\int_t^{t_1}\hat{f}'_x(\tau)d\tau\rangle|_{t_0}^{t_1}+\lambda\int_{t_0}^{t_1}\langle\dot{h}(t),\int_t^{t_1}\hat{f}'_x(\tau)d\tau\rangle dt=$$

$$=\lambda\langle h(t_0),\int_{t_0}^{t_1}\hat{f}'_x(\tau)d\tau\rangle+\lambda\int_{t_0}^{t_1}\langle\dot{h}(t),\int_t^{t_1}\hat{f}'_x(\tau)d\tau\rangle dt \qquad (18)$$

Учитывая в приведенных ниже выкладках, что при любом t $\hat{\varphi}'_x(t)$- $n\times n$ - матрица, а $h(t),d\mu(t)$ - n-мерные вектора, а также используя теорему Фубини о замене порядка интегрирования, получим:

$$\int_{t_0}^{t_1}\langle\hat{\varphi}'_x(t)h(t),d\mu(t)\rangle=\int_{t_0}^{t_1}\langle(h(t_0)+\int_{t_0}^{t}\dot{h}(\tau)d\tau),(\hat{\varphi}'_x(t))^*d\mu(t)\rangle=$$

$$=\int_{t_0}^{t_1}\langle h(t_0),(\hat{\varphi}(t))^*d\mu(t)\rangle+\int_{t_0}^{t_1}\langle\dot{h}(\tau),\int_{\tau}^{t_1}(\hat{\varphi}'_x(t))^*d\mu(t)\rangle d\tau \qquad (19)$$

Подставляя правые части равенств (18),(19) в (17), получим:

$$\langle\hat{\mathcal{L}}'_x,h\rangle=\langle h(t_0),\int_{t_0}^{t_1}\hat{\lambda}f'_x(\tau)d\tau+\int_{t_0}^{t_1}(\hat{\varphi}(t))^*d\mu(t)\rangle+\langle\nu,\hat{g}'_{x(t_0)}h(t_0)\rangle+\langle\nu,\hat{g}'_{x(t_1)}h(t_1)\rangle+$$

$$+\int_{t_0}^{t_1}\langle\dot{h}(t),d\mu(t)\rangle+\int_{t_0}^{t_1}\langle\dot{h}(t),\left(\int_t^{t_1}\lambda\hat{f}'_x(\tau)d\tau+\int_t^{t_1}(\hat{\varphi}'_x(\tau))^*d\mu(\tau)\right)\rangle dt=0$$

В случае, когда $h(t_0)=h(t_1)=0$, это соотношение принимает вид:

$$\langle\hat{\mathcal{L}}'_x,h\rangle=\int_{t_0}^{t_1}\langle\dot{h}(t),d\mu(t)\rangle+\int_{t_0}^{t_1}\langle\dot{h}(t),\left(\int_t^{t_1}\lambda\hat{f}'_x(\tau)d\tau+\int_t^{t_1}(\hat{\varphi}'_x(\tau))^*d\mu(\tau)\right)\rangle dt=0, \qquad (20)$$

и значит

$$d\mu(t)=\left(c-\int_t^{t_1}\lambda\hat{f}'_x(\tau)d\tau-\int_t^{t_1}(\hat{\varphi}'_x(\tau))^*d\mu(\tau)\right)dt \qquad (21)$$

откуда $d\mu(t)=p(t)dt$, где $p(t)$ - ограниченная интегрируемая функция. Подставляя в (21) $p(t)dt$ вместо $d\mu(t)$ и сокращая на dt, получим, что

$$p(t)=\left(c-\int_t^{t_1}\lambda\hat{f}'_x(\tau)d\tau-\int_t^{t_1}(\hat{\varphi}'_x(\tau))^*p(\tau)d\tau\right) \qquad (22)$$

Из (22) следует, что функция $p(t)$ непрерывно дифференцируема. Теперь, заменяя в (20) $d\mu(t)$ на $p(t)dt$ легко вывести (9) с помощью интегрирования по частям. После этого из (17) с $d\mu(t)$, заменённым на $p(t)dt$ и (9) следует (10), а из (16) выводится (11). $\square$

10. Принцип максимума для задач оптимального управления

10.1. **Вводные замечания.** Рассмотрим немного измененную задачу Лагранжа

$$J(x,u)=\int_{t_0}^{t_1} f(t,x(t),u(t))dt\to \min, \tag{1}$$

$$\dot{x}(t)+\varphi(t,x(t),u(t))=0,\quad g(x(t_0),x(t_1))=0, \tag{2}$$

$$u(t)\in U\subset \mathbb{R}^m. \tag{3}$$

Здесь U -замкнутое множество в $\mathbb{R}^m$,

$$f\in C^1([t_0,t_1]\times\mathbb{R}^{n+m}),\quad \varphi\in C^1([t_0,t_1]\times\mathbb{R}^{n+m};\mathbb{R}^n),\quad g\in C^1(\mathbb{R}^{2n};\mathbb{R}^k),$$

-заданные функции, а искомыми являются фазовая функция и управление $x(t),u(t)$.

Примеры показывают, что часто решение $\hat{x},\hat{u}$ этой задачи не принадлежит пространству $C^1\times C$. Более подходящими здесь являются следующие пространства:

Определение 10.1. *Пространство* $KC([t_0,t_1];\mathbb{R}^m)$ *- это множество кусочно-непрерывных функций с конечным числом точек разрыва* $t_0<\tau_1<\dots<\tau_r<t_1$ *первого рода (т.е. в точке разрыва функция непрерывна как слева, так и справа).*

Определение 10.2. *Пространство* $KC^1([t_0,t_1];\mathbb{R}^n)$ *- это множество функций, непрерывных на отрезке* $[t_0,t_1]$*, таких что у каждой первая производная принадлежит* $KC([t_0,t_1];\mathbb{R}^n)$.

Определение 10.3. *Пара* $(\hat{x}(t),\hat{u}(t))\in KC^1([t_0,t_1];\mathbb{R}^n)\times KC([t_0,t_1];\mathbb{R}^m)$ *называется допустимой для задачи (1)-(3), если она удовлетворяет условиям (2),(3). Символом* $\mathcal{A}$ *обозначается множество всех допустимых пар задачи (1)-(3).*

Определение 10.4. *Пара функций* $(\hat{x}(t),\hat{u}(t))\in\mathcal{A}$ *называется оптимальным процессом (точкой минимума) задачи (1)-(3), если при некотором* $\varepsilon>0$ *для любой пары* $(x,u)\in\mathcal{A}$ *такой, что* $\|x-\hat{x}\|_C<\varepsilon$ *справедливо неравенство* $J(\hat{x},\hat{u})\le J(x,u)$.

Отметим также, что так как $(\hat{x}(t),\hat{u}(t))\in KC^1([t_0,t_1];\mathbb{R}^n)\times KC([t_0,t_1];\mathbb{R}^m)$, то такой локальный минимум называется сильным, в отличие от случая $(\hat{x}(t),\hat{u}(t))\in C^1([t_0,t_1];\mathbb{R}^n)\times C([t_0,t_1];\mathbb{R}^m)$, когда локальный минимум называется слабым.

10.2. **Необходимые условия: принцип Лагранжа.** Покажем как для задачи (1)-(3) формулируется принцип Лагранжа. Прежде всего необходимо выписать функцию Лагранжа. Она имеет такой же вид, как и в случае задачи Лагранжа, рассмотренной в предыдущем параграфе:

$$\mathcal{L}(x,u,\lambda,p,\nu)=\int_{t_0}^{t_1}\lambda f(t,x(t),u(t))+\langle p(t),\dot{x}(t)+\varphi(t,x(t),u(t))\rangle dt$$
$$+\langle\nu,g(x(t_0),x(t_1))\rangle.$$

Чтобы вывести необходимые условия экстремума, мы должны изучить две задачи на минимум. Во-первых, минимизацию функции Лагранжа по x, когда в ней вместо u поставлено оптимальное управление $\hat{u}$:

$$1)\mathcal{L}(x, \hat{u}, \lambda, p, \nu) \to \min$$

Эта задача решается как в случае задачи Лагранжа и в результате получаются те же уравнения (9),(10), что и в теореме 9.2.

Совсем другая ситуация возникает при минимизации функции Лагранжа по управлению u, когда в ней вместо x поставлена фазовая компонента $\hat{x}$ решения задачи оптимизации:

$$2)\mathcal{L}(\hat{x}, u, \lambda, p, \nu) \to \min,\ u(t) \in U.$$

Так как здесь функция Лагранжа минимизируется по u, пробегающему замкнутое множество U, производная функции Лагранжа, приравненная к нулю, очевидно, никак не связана с решением. Подход к решению в такой ситуации был предложен Л.С.Понтрягиным в 1953г., и состоит в следующем. Убрав из функции Лагранжа все члены, не содержащие u, перепишем задачу 2) в виде:

$$\int_{t_0}^{t_1} \lambda f(t, \hat{x}(t), u(t)) + \langle p(t), \varphi(t, \hat{x}(t), u(t))\rangle dt \to \min,\ u(t) \in U.$$

Очевидно, этот интеграл достигает минимального значения, если при почти всех $t \in [t_0, t_1]$ минимального значения достигает подинтегральная функция. Следовательно, оптимальное управление $\hat{u}(t)$ определяется из соотношения:

$$\lambda f(t, \hat{x}(t), \hat{u}(t)) + \langle p(t), \varphi(t, \hat{x}(t), \hat{u}(t))\rangle =$$
$$= \min_{v\in U} \lambda f(t, \hat{x}(t), v) + \langle p(t), \varphi(t, \hat{x}(t), v)\rangle,$$

которое обычно называют принципом максимума Понтрягина.

Лекция 13

10.3. **Задача со свободным концом.** Так называется следующая задача:

$$J(x, u) = \int_{t_0}^{t_1} f(t, x(t), u(t))dt \to \min, \tag{1}$$

$$\dot{x}(t) + \varphi(t, x(t), u(t)) = 0, \tag{2}$$

$$x(t_0) = x_0, \tag{3}$$

$$u(t) \in U. \tag{4}$$

Для этой задачи мы приведем доказательство необходимых условий минимума, включающих принцип максимума Понтрягина.

Теорема 10.1. *Пусть $f \in C^1([t_0, t_1]\times\mathbb{R}^{n+m})$, $\varphi \in C^1([t_0, t_1]\times\mathbb{R}^{n+m}; \mathbb{R}^n)$, $\mathbb{R}^n \supset U$ - замкнутое множество, $(\hat{x}(t), \hat{u}(t)) \in KC^1([t_0, t_1]; \mathbb{R}^n) \times KC([t_0, t_1]; \mathbb{R}^m)$ - оптимальный процесс в задаче (1)-(4). 1)Тогда существует такое $p(t) \in KC^1([t_0, t_1]; \mathbb{R}_n)$, что*

$$\dot{p}(t) - (\hat{\varphi}'_x)(t))^* p(t) = \hat{f}_x(t), \tag{5}$$

$$p(t_1) = 0. \tag{6}$$

2) Для всех $t \in [t_0, t_1]$, являющихся точками непрерывности оптимального управления $\hat{u}(t)$, справедлив принцип максимума Понтрягина:

$$f(t, \hat{x}(t), \hat{u}(t)) + \langle p(t), \varphi(t, \hat{x}(t), \hat{u}(t))\rangle =$$

$$= \min_{v\in U}(f(t,\hat x(t),v) + \langle p(t), \varphi(t,\hat x(t),v)\rangle). \tag{7}$$

Доказательство. Обсудим сначала первое утверждение теоремы, т.е. соотношения (5),(6). Если известно, что сильный локальный минимум $(\hat x(t),\hat u(t)) \in C^1([t_0,t_1];\mathbb{R}^n)\times C([t_0,t_1];\mathbb{R}^m)$, то доказательство состоит в дословном повторении доказательства теоремы 9.2 с некоторыми упрощениями, связанными с более простым граничным условием (3) Рассмотрим теперь случай, когда $\hat u(t)$ имеет разрывы первого рода в точках $\hat t_j \in (t_0,t_1), j = 1,\dots,r$, где $\hat t_j < \hat t_{j+1}\ \forall j$. Тогда в силу (2) в тех же точках и только в них будут разрывы первого рода и у функции $\dot{\hat x}(t)$, а в остальных точках она будет непрерывна. Для вывода соотношений (5),(6) будем как обычно исследовать уравнение $\mathcal{L}'_x(\hat x,\hat u,p)[h] = 0$ для производной функции Лагранжа по x из задачи $\mathcal{L}(x,\hat u,p)\to\min$. При этом строить уравнение (5) будем последовательно на каждом интервале непрерывности функции $\hat u(t)$. На первом шаге рассмотрим интервал $(\hat t_r, t_1)$ и будем считать, что вариация $h(t)$ в производной $\mathcal{L}'_x[h]$ равна нулю вне этого интервала, т.е. при $t \le \hat t_r$. Тогда, очевидно,

$$0 = \mathcal{L}'_x(\hat x,\hat u,p)[h] = \lim_{\alpha\to 0}\frac{1}{\alpha}(\mathcal{L}(\hat x+\alpha h,\hat u,p) - \mathcal{L}(\hat x,\hat u,p)) =$$

$$= \int_{\hat t_r}^{t_1}(\langle f'_x(t,\hat x(t),\hat u(t)), h(t)\rangle + \langle p(t), \dot h(t) + \varphi'_x(t,\hat x(t),\hat u(t))h(t)\rangle)dt$$

Применяя к правой части последнего из равенств, т.е. для интеграла по $t \in [\hat t_r, t_1]$, рассуждения из теоремы 9.2 о выборе функционала μ^* и построении функции $p(t)$, получим доказательство разрешимости задачи (5),(6) на отрезке $t \in [\hat t_r, t_1]$. Продолжение решения задачи (5),(6) на отрезок времени $t \in [\hat t_r, t_1]$ делается как и выше, только вместо (6) используется равенство $p(t)|_{t=\hat t_r} = p(\hat t_r)$, причем данное Коши $p(\hat t_r)$ было получено на предыдущем шаге. Продолжение решения задачи (5),(6) на последующие временные отрезки гладкости функций $\hat x(t), \hat u(t)$ проводится точно также.

Перейдем к доказательству основного утверждения теоремы (о принципе максимума Понтрягина). Начнем с определения игольчатой вариации.

Пусть $\tau \in (t_0,t_1)$ - точка непрерывности оптимального управления $\hat u(t)$, $0 < \lambda < \lambda_0$, где λ_0 столь мало, что на интервале $(\tau-\lambda,\tau)$ нет разрывов у $\hat u(t)$. Положим

$$u_\lambda(t) = \begin{cases} v, & t \in (\tau-\lambda,\tau),\\ \hat u(t), & t\in[t_0,t_1]\setminus(\tau-\lambda,\tau), \end{cases} \tag{8}$$

где $v \in U$, (U - множество из соотношения (3)), а через $x_\lambda(t)$ обозначим решение задачи

$$\dot x_\lambda(t) + \varphi(t, x_\lambda(t), u_\lambda(t)) = 0, \quad x_\lambda|_{t=t_0} = x_0. \tag{9}$$

Тогда пара (τ, v), из определения функции (8) называется элементарной иголкой, функция $u_\lambda(t)$ называется игольчатой вариацией оптимального управления $\hat u(t)$, а $x_\lambda(t)$ - игольчатой вариацией оптимальной траектории $\hat x(t)$.

Лемма 10.1. *Пусть элементарная иголка (τ, v) фиксирована. Тогда существует такое $\lambda_0 > 0$, что при $\lambda \in (0,\lambda_0)$*

1)траектория $x_\lambda(\cdot)$ определена на всем отрезке $[t_0,t_1]$ и сходится равномерно к $\hat x(\cdot)$ при $\lambda\to 0$;

2) При $t \geq \tau$ рассмотрим $y(t) = \frac{d}{d\lambda} x_\lambda(t)|_{\lambda=0}$. Тогда $y(t)$ удовлетворяет условиям:

$$\dot{y}(t) + \varphi'_x(t, \hat{x}(t), \hat{u}(t))y(t) = 0,\ y|_{t=\tau} = \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau)) - \varphi(\tau, \hat{x}(\tau), v). \tag{10}$$

Доказательство. При $t, s \in (\tau-\lambda, \tau)$ обозначим через $Y(t, s, \xi)$ решение задачи Коши

$$Y'_t(t, s, \xi) + \varphi(t, Y(t, s, \xi), v) = 0,\ Y|_{t=s} = \xi. \tag{11}$$

Тогда $Y(t, s, \xi) \in C^1([\tau - \lambda_0, \tau]^2 \times \mathbb{R}^n)$, и в силу непрерывной зависимости решений этой задачи от исходных данных, для любого $\varepsilon > 0$ существует такое δ, что при $|t - \hat{t}| < \delta, |s - \hat{s}| < \delta, |\xi - \hat{\xi}| < \delta$ справедливо неравенство $\|Y(t, s, \xi) - Y(\hat{t}, \hat{s}, \hat{\xi})\| < \varepsilon$.

Докажем первое утверждение леммы. а) При $t \in [t_0, \tau - \lambda]$, так как $x_\lambda(t) \equiv \hat{x}(t)$, утверждение очевидно. б) При $t \in (\tau - \lambda, \tau)$, учитывая, что $\hat{x}(\tau - \lambda) = Y(\tau-\lambda, \tau-\lambda, \hat{x}(\tau-\lambda))$, будем иметь $\|x_\lambda(t)-\hat{x}(t)\| \leq \|Y(t, \tau-\lambda, \hat{x}(\tau-\lambda))-Y(\tau-\lambda, \tau - \lambda, \hat{x}(\tau - \lambda))\| + \|\hat{x}(\tau - \lambda) - \hat{x}(t)\| \leq 2\varepsilon$ при достаточно малых λ. Поэтому $x_\lambda(t)$ равномерно стремится к $\hat{x}(t)$ на $(\tau - \lambda, \tau)$ при $\lambda \to 0$. в) При $t > \tau$, в силу того, что $x_\lambda(\tau) \to \hat{x}(\tau)$ при $\lambda \to 0$, функция $x_\lambda(t)$ сходится равномерно к $\hat{x}(t)$ при $t \in (\tau, t_1)$, так как $x_\lambda(t)$ - решение уравнения $\dot{x}_\lambda + \varphi(t, x_\lambda, \hat{u}(t)) = 0$ при $t > \tau$.

Докажем второе утверждение леммы. Пусть $\xi_\lambda = x_\lambda(\tau) = Y(\tau, \tau-\lambda, \hat{x}(\tau-\lambda))$, где функции x_λ, Y определены в (9),(11). Докажем второе из соотношений (10). Справедливы равенства:

$$\frac{d}{d\lambda}\xi_\lambda|_{\lambda=0} = \frac{d}{d\lambda} Y(\tau, \tau - \lambda, \hat{x}(\tau - \lambda))|_{\lambda=0} =$$

$$= -Y'_s(\tau, \tau, \hat{x}(\tau)) + Y'_\xi(\tau, \tau, \hat{x}(\tau))\varphi(\tau, \hat{x}(\tau), \hat{u}(\tau))/ \tag{12}$$

Перепишем (11) в виде интегрального соотношения

$$Y(t, s, \xi) = \xi - \int_s^t \varphi(\sigma, Y(\sigma, s, \xi), v)d\sigma$$

из которого, очевидно, следуют равенства:

$$Y'_s|_{t=s=\tau, \xi=\hat{x}(\tau)} = \left(\varphi(\tau, Y(\tau, s, \xi), v) - \int_s^t \varphi'_x Y'_s d\sigma\right)|_{t=s=\tau, \xi=\hat{x}(\tau)} = \varphi(\tau, \hat{x}(\tau), v)$$

$$Y'_\xi|_{t=s=\tau, \xi=\hat{x}(\tau)} = \left(E - -\int_s^t \varphi'_x Y'_\xi d\sigma\right)|_{t=s=\tau, \xi=\hat{x}(\tau)} = E.$$

Подставляя правые части последних двух равенств в (12), получим справедливость второго равенства из (10). Дифференцируя по λ дифференциальное уравнение (9) при $t > \tau$, учитывая (8), и взяв $\lambda = 0$, получим, что $y(t)$ удовлетворяет дифференциальному уравнению (10). $\square$

В силу определения локального минимума $(\hat{x}(t), \hat{u}(t))$ и игольчатой вариации $(x_\lambda(t), u_\lambda(t))$ справедливо неравенство

$$\sigma(\lambda) = \int_{t_0}^{t_1} (f(t, x_\lambda(t), u_\lambda(t)) - f(t, \hat{x}(t), \hat{u}(t)))dt \geq 0,\ \forall\lambda \in (0, \lambda_0) \tag{13}$$

Лемма 10.2. *Для функции $\sigma(\lambda)$, определенной в (13), справедливо соотношение*

$$\frac{d\sigma(\lambda)}{d\lambda}|_{\lambda=0} = f(\tau, \hat{x}(\tau), v) - f(\tau, \hat{x}(\tau), \hat{u}(\tau)) +$$

$$\langle \varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau)), p(\tau) \rangle \geq 0. \quad (14)$$

Доказательство. Так как при $t < \tau - \lambda$ $x_\lambda(t) = \hat{x}(t), u_\lambda(t) = \hat{u}(t)$, то в интеграле (13) предел интегрирования t_0 можно заменить на $\tau - \lambda$, и поэтому

$$d\sigma(\lambda)/d\lambda|_{\lambda=0} = I_1 + I_2,$$

где

$$I_1 = \lim_{\lambda \to 0} \frac{1}{\lambda} \int_{\tau-\lambda}^{\tau} (f(t, x_\lambda(t), v) - f(t, \hat{x}(t), \hat{u}(t)))dt, \ I_2 = \int_{\tau}^{t_1} \frac{\partial}{\partial \lambda} f(t, x_\lambda(t), \hat{u}(t))dt|_{\lambda=0}.$$

В силу утверждения 1) леммы 10.1

$$I_1 = f(\tau, \hat{x}(\tau), v) - f(\tau, \hat{x}(\tau), \hat{u}(\tau)). \quad (15)$$

Для вычисления интеграла I_2, используя уравнение (5) из теоремы 10.1 и утверждение 2) леммы 10.1 получим:

$$\frac{d}{dt}\langle p(t), y(t) \rangle = \langle \dot{p}, y \rangle + \langle p, \dot{y} \rangle = \langle p(t), \hat{\varphi}'_x(t) y(t) \rangle +$$

$$+\langle \hat{f}'_x(t), y(t) \rangle - \langle p(t), \hat{\varphi}'_x(t) y(t) \rangle = \langle \hat{f}'_x(t), y(t) \rangle$$

Подставляя это равенство в интеграл, определяющий I_2, получим, учитывая (6), что

$$I_2 = \int_{\tau}^{t_1} \langle f'_x(t, x_\lambda(t), \hat{u}(t)), \frac{d}{d\lambda} x_\lambda(t) \rangle dt|_{\lambda=0} = \int_{\tau}^{t_1} \langle \hat{f}'_x(t), y(t) \rangle dt = \quad (16)$$

$$= \int_{\tau}^{t_1} \frac{d}{dt} \langle p(t), y(t) \rangle dt = -\langle p(\tau), y(\tau) \rangle = +\langle \varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau)), p(\tau) \rangle.$$

Складывая равенства (15),(16), получим равенство (14). Из неравенства (13) следует неравенство в (14). □

□

Лекция 14

11. Существование решений экстремальных задач

Этой теме посвящены последние три лекции курса. Для более детального знакомства с этой темой см. книгу (уже упоминавшуюся в Лекции 1) Э.М.Галеев, М.И.Зеликин, С.В.Конягин и др. "Оптимальное управление". Москва, Издательство МЦНМО, 2008

11.1. **Предварительные сведения.** Напомним несколько определений. Пусть X — топологическое пространство[2] и $\overline{\mathbb{R}}$ — расширенная прямая, т.е. $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$, где $\mathbb{R}$ — вещественная ось.

Функция $f : X \to \overline{\mathbb{R}}$ называется *собственной*, если она не равна тождественно $+\infty$, и для всех $x \in X, f(x) > -\infty$.

Функция $f : X \to \overline{\mathbb{R}}$ называется *полунепрерывной снизу на* X, если выполнено одно из условий:

i) Для любого $x \in X$ и любой последовательности $\{x_n\}_{n\in\mathbb{N}}$, сходящейся к x при $n \to \infty$

$$f(x) \leq \liminf_{n\to\infty} f(x_n) := \lim_{n\to\infty} \inf_{i\geq n} f(x_i),$$

ii) Для любого $\lambda \in \mathbb{R}$ Лебегово множество функции f, т.е. $\mathcal{L}_\lambda f := \{x \in X \mid f(x) \leq \lambda\}$, замкнуто.

iii) Надграфик функции f, т.е. множество $\mathrm{epi} f := \{(x, \alpha) \in X \times \mathbb{R} \mid \alpha \geq f(x)\}$, замкнут в $X \times \mathbb{R}$.

Предложение 11.1. *Свойства* i), ii), iii) *эквивалентны.*

Доказательство. i) $\Rightarrow$ ii). Пусть $\lambda \in \mathbb{R}, x_n \in \mathcal{L}_\lambda f$ и $x_n \to \widehat{x}$ при $n \to \infty$. Так как согласно i) $f(\widehat{x}) \leq \liminf f(x_n) \leq \lambda$, то $\widehat{x} \in \mathcal{L}_\lambda f$, что доказывает замкнутость $\mathcal{L}_\lambda f$.

ii) $\Rightarrow$ iii). Пусть $(x_n, \alpha_n) \in \mathrm{epi} f$, т.е. $\alpha_n \geq f(x_n)$, и $x_n \to \widehat{x}, \alpha_n \to \widehat{\alpha}$ при $n \to \infty$. Допустим, что $\widehat{\alpha} < f(\widehat{x})$, т.е. iii) не верно, и положим $\varepsilon = (f(\widehat{x}) - \widehat{\alpha})/2$. Так как $\alpha_n \to \widehat{\alpha}$, то при любом достаточно большом n $\alpha_n \leq \widehat{\alpha} + \varepsilon$. Отсюда и из условия $f(x_n) \leq \alpha_n$ в силу ii) следует, что $f(\widehat{x}) \leq \widehat{\alpha} + \varepsilon$. Подставляя в это неравенство выражение для ε, получим противоречие с предположением, что $f(\widehat{x}) > \widehat{\alpha}$.

iii) $\Rightarrow$ i). Пусть i) не верно, т.е. существует такая последовательность

$$x_n \to \widehat{x} \;\text{ при } n \to \infty, \text{ что } f(\widehat{x}) > \beta := \liminf_{n\in\infty} f(x_n).$$

Следовательно $\beta < \infty$. Пусть $\{x_m\}$ – подпоследовательность последовательности $\{x_n\}$ такая, что $f(x_m) \to \beta$ при $m \to \infty$. Если $\beta > -\infty$, то при любом m

$$(x_m, f(x_m)) \in \mathrm{epi} f, \; (x_m, f(x_m)) \to (\widehat{x}, \beta) \;\text{ при } m \to \infty.$$

Но $(\widehat{x}, \beta) \notin \mathrm{epi} f$, что противоречит iii).

Если $\beta = -\infty$, то для любого $\alpha \in \mathbb{R}$ при достаточно больших m справедливо неравенство $\alpha > f(x_m)$, и следовательно $(x_m, \alpha) \in \mathrm{epi} f$. Конечно, $(x_m, \alpha) \to (\widehat{x}, \alpha)$ при $m \to \infty$. Но если взять $\alpha < f(\widehat{x})$, то $(\widehat{x}, \alpha) \notin \mathrm{epi} f$, что противоречит iii). $\square$

Контрольный вопрос 1.1 (далее КВ): Привести пример разрывной функции, полунепрерывной снизу.

В основе теории существования решения экстремальных задач лежит

[2]Топологическим пространством (X, τ) называется множество X и система τ его подмножеств, содержащих $X, \emptyset$, и замкнутых относительно операции объединения любой совокупности множеств и пересечения любого конечного числа множеств. В топологическом пространстве определено понятие сходимости. Множества из τ называют *открытыми*, а их дополнения *замкнутыми*. Топологическое пространство называется *компактным* (или *компактом*), если из любого его покрытия можно выбрать конечное подпокрытие.

Теорема 11.1. *(Принцип компактности Вейерштрасса — Лебега) Пусть X — компактное топологическое пространство и f — собственная полунепрерывная снизу функция на X. Тогда f ограничена снизу и существует точка $\widehat{x} \in X$, в которой f достигает абсолютного минимума.*

Доказательство. Рассмотрим $\mathcal{L}_n f$, $n \in \mathbb{Z}$. Из определения ii) полунепрерывности снизу следует, что $U_n := X \setminus \mathcal{L}_n f$, $n \in \mathbb{Z}$ открытые множества в X. Ясно, что $\ldots \subset U_n \subset U_{n-1} \subset \ldots$ и что $\{U_n\}_{n\in\mathbb{Z}}$ — открытое покрытие в X. Из определения компактности существует число $m \in \mathbb{Z}$ такое, что $U_m = X$, т. е. f ограничена снизу и потому существует нижняя грань $\mu := \inf_{x\in X} f(x)$. Если нижняя грань не достигается,то положив $V_n := X \setminus \mathcal{L}_{\mu+1/n} f$, $n \in \mathbb{N}$, получим, что $\{V_n\}_{n\in\mathbb{N}}$ есть открытое покрытие X. Следовательно, (снова из-за определения компактности) существует число $s \in \mathbb{N}$ такое, что $X = V_s$, т. е. $f > \mu + 1/s$. Но это противоречит определению μ. $\square$

О том, насколько существенно требование компактности области определения минимизируемой функции свидетельствует пример задачи минимизации $\frac{1}{1+x^2} \to \min$, $x \in \mathbb{R}$, где решения нет.

Отметим, что в задачах вариационного исчисления минимизируются функционалы, которые обычно определены на неограниченных множествах бесконечномерных пространств. Так как такие множества не являются компактными, непосредственное применение теоремы 11.1 в такой ситуации невозможно.

Установим принцип компактности, применимый к задачам вариационного исчисления.

11.2. **Принцип компактности.** Пусть X–нормированное пространство, т.е. линейное пространство, снабженное нормой $\|\cdot\|$. В нем помимо сходимости по норме $\|\cdot\|$ (т.е. сильной сходимости) существует и другой тип предельного перехода. Последовательность элементов $\{x_n\}_{n\in\mathbb{N}}$ нормированного пространства X называется *слабо сходящейся*, если для любого линейного ограниченного функционала x^* на X (т.е. для $x^* \in X^*$)

$$\langle x^*, x_n\rangle \to \langle x^*, x\rangle \qquad n \to \infty$$

где $\langle x^*, x\rangle$ — значение функционала x^* на векторе x.

Подмножество $\mathcal{A}$ пространства X называется *секвенциально слабо замкнутым*, если предел $\widehat{x}$ любой слабо сходящийся последовательности его элементов $x_n \in \mathcal{A}$ принадлежит $\mathcal{A}$.

Кв 1.2. Пусть X– нормированное пространство и A– (сильно) замкнутое поднножество X. Всегда ли оно является секвенциально слабо замкнутым?

Напомним, что нормированное пространство называется *банаховым*, если оно является полным, т.е. любая его фундаментальная последовательность элементов сходится. Банахово пространство X называется *рефлексивным*, если из всякой ограниченной последовательности его элементов можно выбрать слабо сходящуюся подпоследовательность. [3]

Примером рефлексивного пространства является пространство $L_p(\Omega)$ (где Ω — область в пространстве $\mathbb{R}^n$) при $1 < p < \infty$ (см. А.Н Колмогоров, С.В.Фомин "Элементы теории функций и функционального анализа").

---

[3]Отметим, что очень часто рефлексивным называют банахово пространство X, совпадающее со своим вторым сопряженным $X^{**} := (X^*)^*$при каноническом вложении X в X^{**}. Эквивалентность этих двух определений составляет содержание теоремы Эберлейна-Шмульяна (см. К.Иосида "Функциональный анализ"изд.Мир 1967. Приложение к гл.V.)

Кв 1.3. Привести пример нормированного, но не банахова пространства.

Задача 1.1. Доказать, что все конечномерные нормированные пространства рефлексивны.

Задача 1. 2*. Привести пример нерефлексивного банахова пространства.

В полном соответствии с определениями i), ii), iii) функций, полунепрерывных снизу, можно дать следующее определение.

Функционал $f : X \to \overline{\mathbb{R}}$ называется *полунепрерывным снизу на X относительно слабой сходимости*, если выполнено одно из условий:

i') Для любого $x \in X$ и любой последовательности $\{x_n\}$, слабо сходящейся к x при $n \to \infty$

$$f(x) \leq \liminf_{n\to\infty} f(x_n)$$

i') Для любого $\lambda \in \mathbb{R}$ лебегово множество $\mathcal{L}_\lambda f$ секвенциально слабо замкнуто.

ii') Надграфик epi$f$ секвенциально слабо замкнут в $X \times \mathbb{R}$.

Предложение 11.2. *Свойства* i'), ii'), iii') *эквивалентны.*

Задача 1.3 Доказать предложение 11.2 (аналогично предложению11.1).

Пусть X — рефлексивное банахово пространство, $\mathcal{A} \subset X$, f — функционал на X. Рассмотрим экстремальную задачу:

$$f(x) \to \min; \quad x \in \mathcal{A}. \tag{11.1}$$

Решение задачи, т. е. точку, в которой достигается глобальный минимум функционала f, обозначим $\hat{x}$.

Говорят, что задача (11.1) *коэрцитивна*, если для некоторого $\lambda \in \mathbb{R}$ лебегово множество $\mathcal{L}_\lambda(f) = \{x \in \mathcal{A} \mid f(x) \leq \lambda\}$ непусто и ограничено в X. (В случае, когда в (11.1) $\mathcal{A} = X$ коэрцитивным часто называют функционал f, а не задачу (11.1)).

Теорема 11.2. *(О существовании точки минимума) Пусть собственный функционал f определен на рефлексивном банаховом пространстве X и полунепрерывен снизу на X относительно слабой сходимости. Тогда, если множество $\mathcal{A}$ секвенциально слабо замкнуто, а задача (11.1) коэрцитивна, то функционал f ограничен снизу и достигает своего абсолютного минимума на $\mathcal{A}$.*

Доказательство. Пусть

$$\mu = \inf_{x\in\mathcal{A}} f(x)$$

и λ — число из определения коэрцитивности задачи (11.1). Так как $\mathcal{L}_\lambda(f) \neq \emptyset$, то $\lambda \geq \mu$. Если $\lambda = \mu$, то любая точка $x \in \mathcal{L}_\lambda f \neq \emptyset$ является точкой абсолютного минимума $f(x)$ и теорема доказана. Рассмотрим случай $\lambda > \mu$. Согласно определению нижней грани существует такая последовательность $x_n \in \mathcal{A}$, что $f(x_n) \to \mu$. Все члены последовательности $\{x_n\}_{n\in\mathbb{N}}$ за исключением, возможно, конечного числа принадлежат $\mathcal{L}_\lambda(f)$ и поэтому множество $\{\|x_n\|\}$ ограничено. Вследствие рефлексивности X, переходя, если нужно, к подпоследовательности, можно считать, что $x_n \to \hat{x}$ слабо в X. Вследствие секвенциальной слабой замкнутости $\mathcal{A}$ $\hat{x} \in \mathcal{A}$, а благодаря полунепрерывности снизу относительно слабой сходимости $f(\hat{x}) \leq \mu = \lim_{n\to\infty} f(x_n)$. Так как по условию $f(\hat{x}) > -\infty$, то μ конечно и $f(\hat{x}) = \mu$, т. е. $\hat{x}$ — точка абсолютного минимума функционала f. $\square$

У теоремы 11.2 имеются два достаточно трудно проверяемых условия: секвенциально слабой замкнутости множества $\mathcal{A}$ и полунепрерывности снизу функционала f относительно слабой сходимости. В следующем пункте мы приведем легко проверяемые условия, гарантирующие их выполнение.

Лекция 15

11.3. **Теорема Мазура и ее следствия.** Начнем с напоминания определений. Если a, b — точки линейного пространства X, то отрезок $[a, b]$, соединяющий эти точки, определяется формулой

$$[a, b] := \{x \in X|\ x = \alpha a + (1-\alpha)b, \qquad \alpha \in [0, 1]\}.$$

Подмножество A пространства X называется *выпуклым*, если для любых точек $a, b \in A$ отрезок $[a, b]$ принадлежит A.

Задача 1.4. Доказать, что замыкание $\overline{A}$ выпуклого множества A выпукло. (Предполагается, что $A \subset X$ и X - нормированное пространство.)

Если $x_1, \ldots, x_n \in X$, то при любых $\alpha_i \geq 0$ таких, что $\Sigma_{i=1}^n \alpha_i = 1$ вектор

$$x = \sum_{i=1}^n \alpha_i x_i \tag{11.2}$$

называется *выпуклой комбинацией* вектров $x_1, \ldots, n$.

Если A — выпуклое множество и $x_1, \ldots, x_n \in A$, то любая выпуклая комбинация (11.2) принадлежит A. Действительно, если в (11.2) $n = 2$, то утверждение следует из определения выпуклого множества. Пусть утверждение доказано для всех выпуклых комбинаций не более, чем $n-1$ элементов и пусть в (11.2) $|\alpha'| = \alpha_1 + \cdots + \alpha_{n-1} > 0$ (в противном случае $x = x_n \in A$). Тогда по предположению индукции $x' = \sum_{i=1}^{n-1} \frac{\alpha_i}{|\alpha'|} x_i \in A$ и по определению выпуклого множества $x = |\alpha'|x' + \alpha_n x_n \in A$.

Пусть B — подмножество в X. *Овыпуклением* B называется множество всех выпуклых комбинаций элементов B. Овыпукление обозначается $\mathrm{Conv}B$. $\mathrm{Conv}B$ — выпуклое множество, потому, что, если

$$y_1 = \sum_{i=1}^n \alpha_i x_i, \qquad y_2 = \sum_{j=1}^m \beta_j z_j$$

— выпуклые комбинации элементов $x_i \in B$, $i = 1, \ldots n$, $z_j \in B$, $j = 1, \ldots, m$, то при любом $\gamma \in (0, 1)$ элемент

$$\gamma y_1 + (1-\gamma)y_2 = \sum_{i=1}^n \gamma\alpha_i x_i + \sum_{j=1}^m (1-\gamma)\beta_j z_j$$

является выпуклой комбинацией элементов из B и поэтому принадлежит $\mathrm{Conv}B$.

Задача 1.5. Вычислить овыпукление объединения двух гипербол на плоскости: $H_1\{(x_1, x_2)|x_1x_2 = 1, x_i > 0, i = 1, 2\}$ и $H_2 = \{x_1x_2 = -1, x_1 < 0, x_2 > 0\}$.

Теорема 11.3. *(Мазур) Пусть $\widehat{x}$ является слабым пределом последовательности $\{x_n \mid n \in \mathbb{N}\}$ при $n \to \infty$. Тогда существует такая последовательность $y_n = \sum_{j=1}^n \alpha_{j,n} x_j$ выпуклых комбинаций элементов x_k, что $y_n \to \widehat{x}$ сильно при $n \to \infty$.*

Доказательство. Допустим противное. Это означает, что $\widehat{x}$ не принадлежит замыканию $\overline{\mathrm{Conv}\{x_n\}}$ овыпукления счетного множества $\{x_n\}$. Так как замыкание выпуклого множества выпукло, то $\overline{\mathrm{Conv}\{x_n\}}$ — замкнутое выпуклое множество, и $\widehat{x} \notin \overline{\mathrm{Conv}\{x_n\}}$. Поэтому в силу теоремы отделимости существует линейный непрерывный функционал x^* такой, что

$$\sup_{z \in \overline{\mathrm{Conv}\{x_n\}}} \langle x^*, z\rangle < \langle x^*, \widehat{x}\rangle - \varepsilon$$

при некотором $\varepsilon > 0$. Но это неравенство противоречит слабой сходимости последовательности x_n к $\widehat{x}$. $\square$

Важную роль в дальнейшем играют два следствия теоремы Мазура.

Следствие 11.1. *Если множество выпукло и замкнуто, то оно секвенциально слабо замкнуто.*

Доказательство. Пусть $x_n \in A$ и $x_n \to \widehat{x}$ слабо при $n \to \infty$. По теореме Мазура существует такая последовательность y_k выпуклых комбинаций элементов x_n, что $y_k \to \widehat{x}$ сильно при $k \to \infty$. Так как A — выпукло, то $y_k \in A$, а из замкнутости A следует, что $\widehat{x} \in A$. $\square$

Следствие 11.2. *Если функционал $f : X \to \overline{\mathbb{R}}$ является выпуклым и полунепрерывным снизу, то он полунепрерывен снизу относительно слабой сходимости.*

Доказательство. Из условий теоремы и свойства iii) полунепрерывности снизу множество epi$f$ выпукло и замкнуто в $X \times \mathbb{R}$. Поэтому согласно следствию 11.1 epi$f$ секвенциально слабо замкнуто, а в силу свойства iii') функция f полунепрерывна снизу относительно слабой сходимости. $\square$

11.4. **Пространства Соболева.** Пусть Ω — область в пространстве $\mathbb{R}^d$, граница которой $\partial\Omega$ является бесконечно дифференцируемым многообразием. Напомним, что символом $C^\infty(\overline{\Omega})$ обозначается пространство бесконечно дифференцируемых функций, определенных на замыкании $\overline{\Omega}$ области Ω, а $C_0^\infty(\Omega)$– это подпространство пространства $C^\infty(\overline{\Omega})$, состоящее из функций с компактным носителем. При этом носитель f (supp$f$) определяется формулой: $\mathrm{supp}f = \overline{\{x \in \Omega : f(x) \neq 0\}}$, причем черта наверху опять обозначает операцию замыкания множества.

Рассмотрим функцию $x \to u(x)$, принадлежащую пространству $L_p(\Omega)$. Ее обобщенной производной $\partial u/\partial x_k$ называется такая обобщенная функция (т. е. линейный непрерывный функционал на пространстве $C_0^\infty(\Omega)$, что

$$\int_\Omega u(x)\frac{\partial \varphi(x)}{\partial x_k}\,dx = -\left\langle \frac{\partial u}{\partial x_k}, \varphi\right\rangle, \quad \forall \varphi \in C_0^\infty(\Omega),$$

причем справа стоит значение обобщенной функции $\frac{\partial u}{\partial x_k}$ на пробной функции φ). Пространством Соболева $W_p^1(\Omega)$, $1 \le p < \infty$, называется множество таких функций $u(\cdot) \in L_p(\Omega)$ у которых все обобщенные производные $\frac{\partial u}{\partial x_k}$, $k = 1, \ldots, d$, принадлежат пространству $L_p(\Omega)$.

Норма в пространстве $W_p^1(\Omega)$ определяется формулой:

$$\|u(\cdot)\|^p_{W_p^1(\Omega)} = \int_\Omega \left(|u(x)|^p + \sum_{k=1}^d \left|\frac{\partial u}{\partial x_k}\right|^p\right)dx = \int_\Omega (|u(x)|^p + |\nabla u(x)|^p)dx.$$

Как и в случае одного переменного, чтобы выписанное выражение действительно определяло норму, элементом пространства нужно считать не фиксированную функцию $u(x)$, а класс эквивалентных функций, совпадающих с $u(x)$ при почти всех $x \in \Omega$). Именно так мы и будем считать, хотя и будем допускать вольность речи, называя элементом пространства Соболева конкретные функции.

Теорема 11.4. *(о полноте) Пространство $W_p^1(\Omega)$ полно.*

Доказательство. Пусть последовательность $\{u_n\}$ фундаментальна в $W_p^1(\Omega)$. Тогда последовательности $\{u_n\}$, $\{\partial u_n/\partial x_j\}$, $j = 1, \dots, d$ фундаментальны в $L_p(\Omega)$, а в силу полноты $L_p(\Omega)$ существуют такие функции u, $u^{(j)}$, $j = 1, \dots, d$, что при $n \to \infty$

$$u_n \to u, \qquad \frac{\partial u_n}{\partial x_j} \to u^{(j)}, \qquad j = 1, \dots, d \quad \text{в} \quad L_p(\Omega)$$

Поэтому при любом $\varphi \in C_0^\infty(\Omega)$

$$\int_\Omega u \frac{\partial \varphi}{\partial x_j}\, dx \overset{n\to\infty}{\longleftarrow} \int_\Omega u_n \frac{\partial \varphi}{\partial x_j}\, dx = -\int_\Omega \frac{\partial u_n}{\partial x_j} \varphi\, dx \overset{n\to\infty}{\longrightarrow} -\int_\Omega u^{(j)} \varphi\, dx.$$

Значит $u^{(j)} = \frac{\partial u}{\partial x_j}$ (в смысле обобщенных функций), что доказывает полноту пространства $W_p^1(\Omega)$. □

Приведем без доказательства [4] следующие фундаментальные утверждения:

Теорема 11.5. *(о плотности для пространства $W_p^1(\Omega)$) Пространство $C^\infty(\overline{\Omega})$ плотно в $W_p^1(\Omega)$ при $1 \le p < \infty$.*

Лемма 11.1. *Пусть область Ω ограничена. Тогда пространство $W_q^1(\Omega)$ вполне непрерывно вложено в $L_q(\Omega)$: $W_q^1(\Omega) \Subset L_q(\Omega)$.*

Если $u(\cdot) \in C^\infty(\overline{\Omega})$, то, очевидно, определено сужение (след) функции $u(\cdot)$ на границу $\partial\Omega$: $\gamma u = u(\cdot)|_{\partial\Omega}$. Имеет место

Теорема 11.6. *(о следе) Оператор следа продолжается с $C^\infty(\bar{\Omega})$ до непрерывного оператора $\gamma : W_p^1(\Omega) \to L_p(\partial\Omega)$, $1 \le p < \infty$.*

Схема доказательства. Сначала для произвольной функции $u(\cdot) \in C^\infty(\overline{\Omega})$ устанавливается оценка

$$\|\gamma u(\cdot)\|_{L_p(\partial\Omega)} \le C \|u(\cdot)\|_{W_p^1(\Omega)}, \tag{11.3}$$

где константа C не зависит от $u(\cdot)$ (и это является главной частью доказательства). Потом для произвольной функции $u(\cdot) \in W_p^1(\Omega)$ с помощью теоремы о плотности выбирается последовательность $u_n(\cdot) \in C^\infty(\Omega)$, сходящаяся к $u(\cdot)$ в $W_p^1(\Omega)$. Эта последовательность фундаментальна в $W_p^1(\Omega)$, а в силу приведенной выше оценки последовательность $\gamma u_n(\cdot)$ фундаментальна в $L_p(\partial\Omega)$. Значит, $\gamma u_n(\cdot)$ имеет предел в $L_p(\partial\Omega)$, который обозначается $\gamma u(\cdot)$. Далее устанавливается, что след $\gamma u(\cdot)$ связан лишь с $u(\cdot)$ (т. е. не зависит от аппроксимирующей последовательности) и для $\gamma u(\cdot)$ и $u(\cdot)$ справедлива приведенная выше оценка. □

[4] Относительно доказательства этой и других теорем о пространствах Соболева см., например, Л.К.Эванс. Уравнения с частными производными. Новосибирск, Тамара Рожковская, 2003.

Такая схема доказательства характерна для теории пространств Соболева: сначала доказываемый факт устанавливается для гладких функций, а потом проводится процесс замыкания.

Задача 3.1. Докажите справедливость неравенства (11.3) для любой фкнкции $u(x) \in C^1(\overline{\Omega})$ если а) $\Omega = (a, b)$ (одномерный случай), б) $\Omega \in \mathbb{R}^d$ - ограниченная область с границей $\partial\Omega \subset C^\infty$.

Определим следующее подпространство пространства $W_p^1(\Omega)$:

$$\overset{o}{W}{}_p^1(\Omega) = \{u(\cdot) \in W_p^1(\Omega) \mid \gamma u(\cdot) = 0\},$$

где γ — оператор сужения на границу $\partial\Omega$. Справедлив следующий аналог теоремы о плотности, который мы также приводим без доказательства.

Теорема 11.7. *(о плотности для пространства $\overset{o}{W}{}_p^1(\Omega)$) Пространство $C_0^\infty(\Omega)$ плотно в $\overset{o}{W}{}_p^1(\Omega)$ при $1 \le p < \infty$.*

В дальнейшем при $p = 2$ мы будем использовать также обозначения: $W_2^1(\Omega) = H^1(\Omega)$, $\overset{o}{W}{}_2^1(\Omega) = H_0^1(\Omega)$.

Наконец, приведем еще следующую важную оценку (оставив ее также без доказательства).

Теорема 11.8. *(неравенство Фридрихса) Для любой функции $u(\cdot) \in \overset{o}{W}{}_p^1(\Omega)$ справедливо неравенство*

$$\int_\Omega |u(x)|^p dx \le c \int_\Omega |\nabla u(x)|^p dx$$

с константой c, зависящей лишь от области Ω.

В силу неравенства Фридрихса норму в пространстве $\overset{o}{W}{}_p^1(\Omega)$ естественно задавать равенством:

$$\|u\|_{\overset{o}{W}{}_p^1(\Omega)} = \left(\int_\Omega |\nabla u(x)|^p dx \right)^{1/p}.$$

Докажем теперь важную для нас теорему о рефлексивности пространств $W_p^1(\Omega)$.

Теорема 11.9. *Пространство $W_p^1(\Omega)$ рефлексивно, если $1 < p < \infty$.*

Доказательство. Напомним, что $L_p(\Omega)$ - рефлексивное пространство, причем его сопряженное можно отождествить с $L_q(\Omega)$ где $\frac{1}{q} + \frac{1}{p} = 1$. Из теоремы о виде функционала на прямом произведении пространств сразу следует, что $(L_p(\Omega))^n := L_p(\Omega) \times \cdots \times L_p(\Omega)$ (n раз) имеет в качестве сопряженного пространство $(L_q(\Omega))^n$, откуда сразу следует его рефлексивность, так как его второе сопряженное совпадает с $(L_p(\Omega))^n$, т.е. с ним же.

Рассмотрим линейный непрерывный оператор $A : W_p^1(\Omega) \to (L_p(\Omega))^{d+1}$, определяемый формулой $Au = (u, \partial u/\partial x_1, \ldots, \partial u/\partial x_d)$. Очевидно, $AW_p^1(\Omega)$ - замкнутое подпространство в $(L_p(\Omega))^{d+1}$ (доказательство такое же как в теореме 11.4), а в силу следствия 11.1 оно секвенциально слабо замкнуто.

Покажем, что для любого функционала $F \in (W_p^1(\Omega))^*$ существует $(f_0, \dots, f_d) \in (L_q(\Omega))^{d+1}$, где $1/p + 1/q = 1$ такой, что

$$F(u) = \int_\Omega (f_0(x)u(x) + \sum_{j=1}^d f_j(x)\frac{\partial u(x)}{\partial x_j})\,dx \quad \forall\, u \in W_p^1(\Omega). \tag{11.4}$$

Действительно, F определяет функционал F_1 на $AW_p^1(\Omega)$. По теореме Хана-Банаха F_1 можно продолжить с $AW_p^1(\Omega)$ на все пространство $(L_p(\Omega))^{d+1}$ с сохранением нормы. Так как функционал на $(L_p(\Omega))^{d+1}$ определяется некоторым элементом $(f_0, \dots, f_d) \in (L_q(\Omega))^{d+1}$, то отсюда следует (11.4).

Пусть последовательность $\{u_n\}$ ограничена в $W_p^1(\Omega)$. Значит последовательность $\{(u_n, \partial u_n/\partial x_j, j = 1, \dots, d)\}$ ограничена в $(L_p(\Omega))^{d+1}$, и в силу рефлексивности этого пространства существует подпоследовательность $\{w_k\} \equiv \{(u_{n_k}, \partial u_{n_k}/\partial x_j, j = 1, \dots, d)\}$ слабо сходящаяся к $\widehat{w} \in (L_p(\Omega))^{d+1}$ Так как $w_k \in AW_p^1(\Omega)$, и множество $AW_p^1(\Omega)$ секвенциально слабо замкнуто, то $\widehat{w}$ имеет вид $\widehat{w} = (\widehat{u}, \partial\widehat{u}/\partial x_j, j = 1, \dots, d)\}$. Значит ввиду (11.4) $u_{n_k} \to \widehat{u}$ слабо в $W_p^1(\Omega)$. $\square$

Задача 3.2. Пусть Y- замкнутое подпространство рефлексивного банахова пространства X. Доказать, что Y- рефлексивное пространство.

Задача 3.3. Доказать рефлексивность пространства $\overset{o}{W}{}_p^1(\Omega)$, если $1 < p < \infty$.

Лекция 16

11.5. **Вариационная задача: коэрцитивность и условия роста.** Мы изучим следующий многомерный аналог простейшей задачи вариационного исчисления:

$$J(y(\cdot)) = \int_\Omega L(x, y(x), \nabla y(x))\,dx \to \min \tag{11.5}$$

при условии, что

$$y(x)|_{x\in\partial\Omega} = 0 \tag{11.6}$$

где $\Omega \subset \mathbb{R}^d$ — ограниченная область с границей $\partial\Omega$ класса C^∞, $x = (x_1, \dots, x_d) \in \Omega$ — независимые переменные, $y(x)$ — искомая вещественнозначная функция, $\nabla y(x) = (\partial y/\partial x_1, \dots, \partial y/\partial x_d)$ — ее градиент. Об интегранте

$$L(x, y, p) = L(x, y, p_1, \dots, p_d)$$

будем предполагать, что

$$L(\cdot,\cdot,\cdot) \in C(\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^d), \quad \frac{\partial L}{\partial p_i}(x, y, p) \in C(\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^d), \quad i = 1, \dots, d. \tag{11.7}$$

Теорема существования решения задачи (11.5), (11.6) (теорема Тонелли) будет доказана с помощью теоремы 11.2 о существовании решения абстрактной задачи. Чтобы гарантировать справедливость условия *коэрцитивности* функционала $J(y)$ из (11.5) мы наложим на функцию $L(x, y, p)$ следующее *условие роста*:

Пусть задано

$$1 < q < \infty \tag{11.8}$$

Существуют константы $\alpha > 0$, $\beta \geq 0$ такие, что неравенство

$$L(x, y, p) \geq \alpha|p|^q - \beta \tag{11.9}$$

справедливо при всех $x \in \overline{\Omega}$, $y \in \mathbb{R}$, $p \in \mathbb{R}^d$.

Если функция L удовлетворяет условию (11.8), (11.9), то задачу (11.5), (11.6) естественно рассматривать на пространстве Соболева $W_q^1(\Omega)$, а ограничение (11.6) записать в виде $y \in \mathcal{A} \equiv \overset{o}{W}{}_q^1(\Omega)$.

Лемма 11.2. *Функционал $J(y)$, $y \in \overset{o}{W}{}_q^1(\Omega)$ ограничен снизу, не равен тождественно $+\infty$ и удовлетворяет условию коэрцитивности.*

Доказательство. Подставляя в (11.9) $(y,p) \to (y(x), \nabla y(x))$ и интегрируя по Ω получим, что

$$J(y) \geq \alpha \int_\Omega |\nabla y(x)|^q dx - \beta \int_\Omega dx = \alpha \|y\|^q_{\overset{o}{W}{}_q^1(\Omega)} - \beta|\Omega|, \tag{11.10}$$

откуда следует ограниченность снизу функционала $J(y)$ на $\overset{o}{W}{}_q^1(\Omega)$.

Так как $L(x,y,p)$ удовлетворяет (11.7), то при $y \in C^1(\bar{\Omega})$, $y|_{\partial\Omega} = 0$ имеем $J(y) < \infty$.

Возьмем произвольную функцию $y_0(x) \in C^1(\bar{\Omega})$, $y_0|_{\partial\Omega} = 0$ и положим $R = J(y_0)$. Очевидно, $\mathcal{L}_R J = \{y \in \overset{o}{W}{}_q^1(\Omega) \colon J(y) \leq R\} \neq \emptyset$. Кроме того, в силу неравенства (11.10)

$$\mathcal{L}_R J \subset \{y \in \overset{o}{W}{}_q^1(\Omega) : \|y\|^q_{\overset{o}{W}{}_q^1(\Omega)} \leq \frac{1}{\alpha}(\beta|\Omega| + R)\}.$$

Следовательно, множество $\mathcal{L}_R J$ ограничено, и функционал J удовлетворяет условию коэрцитивности. $\square$

Итак, в лемме 11.2 показано, что условия (11.8), (11.9) гарантируют выполнение условия коэрцитивности у задачи (11.5), (11.6).

11.6. **Квазирегулярность и полунепрерывность снизу относительно слабой сходимости.** Запись задачи (11.5), (11.6) в форме абстрактной задачи (11.1) такова:

$$J(y) \to \min, \quad y \in \overset{o}{W}{}_q^1(\Omega), \tag{11.11}$$

причем функционал J задается равенством (11.5) и определен на $W_q^1(\Omega)$. Будучи замкнутым линейным подпространством пространства $W_q^1(\Omega)$, множество $\overset{o}{W}{}_q^1(\Omega)$ является секвенциально слабо замкнутым. Таким образом, для задачи (11.11) проверены все условия абстрактной теоремы существования, за исключением условия полунепрерывности снизу относительно слабой сходимости для функционала J. Оказывается, чтобы обеспечить выполнимость этого условия для интегрального функционала вида (11.5) не обязательно требовать выпуклости этого функционала, а достаточно требовать выпуклости интегранта $L(x,y,p)$ только по переменным p, т. е. *его квазирегулярности.* Наложим это условие:

$$\text{функция } p \mapsto L(x,y,p) \ \text{ выпукла при любых } x \in \bar{\Omega},\ y \in \mathbb{R}. \tag{11.12}$$

Теорема 11.10. *Пусть интегрант L ограничен снизу и удовлетворяет условиям квазирегулярности* (11.12) *и гладкости* (11.7). *Тогда функционал J, определенный в* (11.5), *полунепрерывен снизу относительно слабой сходимости на $W_q^1(\Omega)$.*

Доказательство. Мы должны показать, что если

$$y_k \to \hat{y} \ \text{ слабо в } W_q^1(\Omega), \tag{11.13}$$

то

$$J(\hat{y}) \le \hat{J} \equiv \liminf_{k\to\infty} J(y_k).$$

Переходя, если нужно, к подпоследовательности, можно считать, что

$$\lim_{k\to\infty} J(y_k) = \hat{J}.$$

В силу (11.13) $\max\limits_k \|y_k\|_{W_q^1} < \infty$ (см.[КФ], гл. 4 § 3) и поэтому, вследствие леммы 11.1 о полной непрерывности вложения $W_q^1(\Omega) \Subset L_q(\Omega)$, $y_k \to \hat{y}$ сильно в $L_q(\Omega)$. Поэтому, переходя если нужно к подпоследовательности, можно считать, что

$$y_k(x) \to \hat{y}(x) \ \text{ при почти всех } x \in \Omega. \tag{11.14}$$

Для любого $\varepsilon > 0$ согласно теореме Егорова ([КФ], гл. 5 § 4)

$$y_k(x) \to \hat{y}(x) \ \text{ равномерно по } x \in E_\varepsilon, \tag{11.15}$$

где E_ε — некоторое измеримое подмножество Ω, для которого

$$\mathrm{mes}(\Omega \setminus E_\varepsilon) \le \varepsilon$$

($\mathrm{mes} A$ — это лебегова мера множества A).

Пусть $F_\varepsilon = \{x \in \Omega\colon |\hat{y}(x)| + |\nabla\hat{y}(x)| \le 1/\varepsilon\}$. Так как $\mathrm{mes}(\Omega \setminus F_\varepsilon) \to 0$ при $\varepsilon \to 0$, то для множества

$$G_\varepsilon = F_\varepsilon \cap E_\varepsilon \tag{11.16}$$

справедливо соотношение

$$\mathrm{mes}(\Omega \setminus G_\varepsilon) \to 0 \ \ \varepsilon \to 0. \tag{11.17}$$

Вследствие (11.13)

$$\int (\varphi(x), \nabla y_k(x) - \nabla\hat{y}(x))\, dx \to 0 \ \text{ при } k \to \infty$$

для любой $\varphi \in (L_r(\Omega))^d \equiv L_r(\Omega) \times \cdots \times L_r(\Omega)$ (d раз) с $\frac{1}{r} + \frac{1}{q} = 1$ и поэтому при любом достаточно малом $\varepsilon > 0$

$$\gamma_k^\varepsilon := \int_{G_\varepsilon} (L_p(x, \hat{y}(x), \nabla\hat{y}(x)), \nabla y_k(x) - \nabla\hat{y}(x))\, dx \to 0 \ \ k \to \infty.$$

В силу (11.14), (11.16) при каждом достаточно малом $\varepsilon > 0$

$$\begin{aligned} \alpha_k^\varepsilon := &\int_{G_\varepsilon} (L(x, y_k(x), \nabla\hat{y}(x)) - L(x, \hat{y}(x), \nabla\hat{y}(x)))\, dx + \\ &+ \int_{G_\varepsilon} (L_p(x, y_k(x), \nabla\hat{y}(x)) - L_p(x, \hat{y}(x), \nabla\hat{y}(x)), \nabla y_k(x) - \nabla\hat{y}(x))\, dx \to 0 \\ &\hspace{30em} k \to \infty \end{aligned}$$

Вследствие (11.12)

$$\begin{gathered} L(x, y_k(x), \nabla y_k(x)) \ge L(x, \hat{y}(x), \nabla\hat{y}(x)) + \\ + [L(x, y_k(x), \nabla\hat{y}(x)) - L(x, \hat{y}(x), \nabla\hat{y}(x))] + \\ + (L_p(x, \hat{y}(x), \nabla\hat{y}(x)), \nabla y_k(x) - \nabla\hat{y}(x)) + \\ + (L_p(x, y_k(x), \nabla\hat{y}(x)) - L_p(x, \hat{y}(x), \nabla\hat{y}(x)), \nabla y_k(x) - \nabla\hat{y}(x)). \end{gathered}$$

Интегрируя это неравенство по G_ε, получим, учитывая определение γ_k^ε, α_k^ε:

$$\int_{G_\varepsilon} L(x, y_k(x), \nabla y_k(x))\, dx \geq \int_{G_\varepsilon} L(x, \hat{y}(x), \nabla \hat{y}(x))\, dx + \gamma_k^\varepsilon + \alpha_k^\varepsilon. \tag{11.18}$$

В соответствии с условием ограниченности снизу интегранта L

$$L(x, y, p) \geq -\beta \quad \forall (x, y, p) \in \bar{\Omega} \times \mathbb{R} \times \mathbb{R}^d$$

с некоторым $\beta \geq 0$. Поэтому в силу (11.17), (11.18) для любого $\delta > 0$ существуют такие $\varepsilon > 0$ и $k_0 > 0$, что при всех $k > k_0$

$$J(u_k) = \int_{G_\varepsilon} L(x, y_k, \nabla y_k)\, dx + \int_{\Omega \setminus G_\varepsilon} L(x, y_k, \nabla y_k)\, dx \geq$$
$$\geq -\beta \operatorname{mes}(\Omega \setminus G_\varepsilon) + J(\hat{y}) - \int_{\Omega \setminus G_\varepsilon} L(x, \hat{y}, \nabla \hat{y})\, dx + \gamma_k^\varepsilon + \alpha_k \geq J(\hat{y}) - \delta.$$

Так как δ здесь произвольно, то $J(\hat{y}) \leq \hat{J} = \lim_{k\to\infty} J(y_k)$. $\square$

11.7. **Теорема Тоннели.** Теперь мы можем доказать основную теорему о существовании решения задачи (11.5), (11.6).

Теорема 11.11. *(Тонелли) Пусть интегрант* $L(x, y, p) \in C(\overline{\Omega} \times \mathbb{R} \times \mathbb{R}^d)$ *из* (11.5) *удовлетворяет условиям гладкости* (11.7)*, роста* (11.8)*,* (11.9) *и выпуклости по* p (11.12)*. Тогда в пространстве* $W_q^1(\Omega)$ *существует решение* $\hat{y}(x) \in W_q^1(\Omega)$ *задачи* (11.5)*,* (11.6)*.*

Доказательство. В силу леммы 11.2 и теоремы 11.10 задача (11.5), (11.6) удовлетворяет всем предположениям теоремы 11.2 о разрешимости абстрактной экстремальной задачи. Поэтому из этой теоремы следует справедливость теоремы Тонелли. $\square$