

Конспект лекций проф. М.И. Зеликина по Вариационному исчислению и оптимальному управлению.

Содержание

1	ПОСТАНОВКА ПРОСТЕЙШЕЙ ЗАДАЧИ КЛАССИЧЕСКОГО ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ	2
2	УРАВНЕНИЕ ЭЙЛЕРА	3
3	ГЕОДЕЗИЧЕСКИЕ НА РИМАНОВОМ МНОГООБРАЗИИ	5
4	ВАРИАЦИЯ ФУНКЦИОНАЛА С ПОДВИЖНЫМИ КОНЦАМИ	8
5	УСЛОВИЯ ТРАНСВЕРСАЛЬНОСТИ	9
6	ПРЕОБРАЗОВАНИЕ ЛЕЖАНДРА	11
7	КАНОНИЧЕСКИЕ ПЕРЕМЕННЫЕ	12
8	МЕХАНИЧЕСКИЙ СМЫСЛ КАНОНИЧЕСКИХ ПЕРЕМЕННЫХ	14
9	УСЛОВИЕ ЛЕЖАНДРА	18
10	СОПРЯЖЕННЫЕ ТОЧКИ	19
11	НЕОТРИЦАТЕЛЬНАЯ ОПРЕДЕЛЕННОСТЬ ВТОРОЙ ВАРИАЦИИ	20
12	УРАВНЕНИЕ РИККАТИ	21
13	ГРАССМАНОВЫ МНОГООБРАЗИЯ	22
14	ТЕОРЕМА ЯКОБИ ОБ ОГИБАЮЩЕЙ	26
15	НЕОБХОДИМОЕ УСЛОВИЕ ВЕЙЕРШТРАССА	30
16	ИНТЕГРАЛЬНЫЙ ИНВАРИАНТ ПУАНКАРЕ-КАРТАНА	32
17	ОСНОВНАЯ КОНСТРУКЦИЯ	34
18	ПОЛЕ ЭКСТРЕМАЛЕЙ	37
19	ЭЙЛЕРОВЫ ЭЛАСТИКИ	38
20	МИНИМАЛЬНЫЕ ПОВЕРХНОСТИ	44

21 ПРАВИЛО МНОЖИТЕЛЕЙ ЛАГРАНЖА ДЛЯ ГЛАДКИХ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ	47
22 ПРИНЦИП МАКСИМУМА ПОНТРЯГИНА	51
23 ДОКАЗАТЕЛЬСТВО ПРИНЦИПА МАКСИМУМА ДЛЯ ЗАДАЧИ СО СВОБОДНЫМ КОНЦОМ	52
24 СУЩЕСТВОВАНИЕ РЕШЕНИЯ ЗАДАЧИ ОПТИМАЛЬНОГО БЫСТРОДЕЙСТВИЯ	54

1 ПОСТАНОВКА ПРОСТЕЙШЕЙ ЗАДАЧИ КЛАССИЧЕСКОГО ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ

В 1696 г. И. Бернулли поставил перед математическим миром задачу о брахистохроне, обещая "воздать заслуженную хвалу тому, кто справится с ее решением". Современники не остались равнодушными и за короткое время Бернулли получил три письма (одно из них анонимное), содержащие различные решения этой задачи. Именно поэтому 1696 год принято считать годом рождения вариационного исчисления. Мы предлагаем читателям последовать примеру анонимного автора (которым был И. Ньютон) и попробовать откликнуться на призыв из далекого 1696 года. Задача о брахистохроне ставилась следующим образом: "В вертикальной плоскости даны две точки a , и b . Определить путь, спускаясь по которому под действием собственной тяжести, тело M , начав двигаться из точки a , дойдет до точки b в кратчайшее время". (Предполагается, что трение отсутствует). Введем в данной плоскости систему координат, приняв за начало координат точку a и направив ось Oy вниз. Пусть $y = y(x)$ – уравнение искомой кривой. По закону сохранения энергии скорость при прохождении точки $(x, y(x))$ будет такой же, как при свободном падении с высоты $y(x)$, т. е. $v = \sqrt{2gy(x)}$. Поэтому, интегрируя вдоль кривой $y = y(x)$, получим, что время движения равно

$$\int \frac{ds}{v} = \int_0^x \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{2gy(x)}} dx. \quad (1)$$

Итак, задача заключается в том, чтобы выбрать функцию $y = y(x)$ удовлетворяющую условиям $y(0) = 0$, $y(x_0) = y_0$ так, чтобы функционал (1) принял наименьшее значение.

Чтобы найти решение некоторой конкретной задачи, математики, как правило, строят общую теорию для решения произвольных задач того же типа. Иногда, после того, как такая теория построена, выясняется, что исходная задача не удовлетворяет тем предположениям, которые были положены в основу теории. Тогда начинается процесс обобщения, который нередко превращает простую исходную конструкцию в существенно более громоздкую. Поскольку нашей целью является скорее построение общей теории, нежели решение задачи о брахистохроне, мы смело пойдем этим путем и рассмотрим при удобных для нас предположениях общую задачу о минимизации интегрального функционала.

Задача 5. Среди всех кривых $x(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n)$, удовлетворяющих граничным условиям

$$x(t_0) = a, \quad x(t_1) = b, \quad (2)$$

где a, b – две заданные точки из \mathbb{R}^n , найти такую, которая дает минимальное значение функционалу

$$\mathcal{J}(x(\cdot)) = \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t)) dt, \quad (3)$$

где $f \in C^2([t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n)$.

Под минимальным значением здесь и далее в гл. V-VII понимается локальный минимум в пространстве $C^1([t_0, t_1])$.

2 УРАВНЕНИЕ ЭЙЛЕРА

Подпространство пространства $C^1([t_0, t_1], \mathbb{R}^n)$, определяемое условиями

$$h(t_0) = h(t_1) = 0 \quad (4)$$

будем обозначать через $\mathcal{C}([t_0, t_1])$ или просто \mathcal{C} .

Пусть кривая $\hat{x}(\cdot)$ есть решение задачи 5. Рассмотрим одно параметрическое семейство кривых $\hat{x}(t) + \lambda h(t)$, где $h(\cdot) \in \mathcal{C}([t_0, t_1])$; $\lambda \in \mathbb{R}$. В силу условий (4) каждая из кривых этого семейства удовлетворяет условиям (2). Рассмотрим скалярную функцию $\varphi(\lambda)$, получаемую подстановкой кривых нашего семейства в функционал (3):

$$\varphi(\lambda) = \int_{t_0}^{t_1} f(t, \hat{x}(t) + \lambda h(t), \dot{\hat{x}}(t) + \lambda \dot{h}(t)) dt. \quad (5)$$

Поскольку $\varphi(0)$ есть значение функционала (3) на $\hat{x}(\cdot)$, функция (5) достигает локального минимума при $\lambda = 0$. Заметим, что из теоремы о дифференцируемости интеграла по параметру и условия $f \in C^2$, ([4] т.2, с.107) следует, что функция $\varphi(\lambda)$ является дважды дифференцируемой по λ . Поэтому $\varphi'(0) = 0$, $\varphi''(0) \geq 0$. (Второе из этих условий будет использовано дальше в гл. VII). Имеем

$$\varphi'(0) = \int_{t_0}^{t_1} [f_x(t, \hat{x}(t), \dot{\hat{x}}(t))h(t) + f_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))\dot{h}(t)] dt = 0. \quad (6)$$

Отметим, что уравнение (6) выполнено для любой функции $h(\cdot) \in \mathcal{C}$.

Лемма (Дюбуа-Реймона).

Пусть $a(t), b(t) \in C([t_0, t_1], \mathbb{R}^n)$ и для любой функции $h(t) \in \mathcal{C}([t_0, t_1])$ выполнено условие

$$\int_{t_0}^{t_1} [a(t)h(t) + b(t)\dot{h}(t)] dt = 0. \quad (7)$$

Тогда функция $b(t)$ непрерывно дифференцируема и

$$-\frac{d}{dt}b(t) + a(t) = 0. \quad (8)$$

Доказательство. Обозначим через $A(t)$ произвольную первообразную функции $a(t)$, т. е. $A(t) = \int_{t_0}^{t_1} a(\tau) d\tau + K$. Интегрируя первое слагаемое в (7) по частям и учитывая (4), получим

$$\int_{t_0}^{t_1} [-A(t) + b(t)] \dot{h}(t) dt = 0. \quad (9)$$

(при любом выборе константы K). Подберем теперь функцию $h(t)$ так, чтобы под интегралом стоял полный квадрат. Для этого положим

$$\int_{t_0}^{t_1} [-A(\tau) + b(\tau)] d\tau. \quad (10)$$

При таком выборе $h(t_0) = 0$, а выполнение условия $h(t_1) = 0$ обеспечивается выбором константы K . Подставив (10) в (9), получим:

$$\int_{t_0}^{t_1} [-A(t) + b(t)]^2 dt = 0,$$

т. е. $A(t) \equiv b(t)$. Следовательно, функция $b(t)$ непрерывно дифференцируема, и выполнено уравнение (8).

Применяя лемму Дюбуа Реймона к (6) получаем, что решение $\hat{x}(t)$ задачи 5 должно удовлетворять системе дифференциальных уравнений

$$-\frac{d}{dt} f_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) + f_x(t, \hat{x}(t), \dot{\hat{x}}(t)) = 0. \quad (11)$$

которая называется уравнением Эйлера. Уравнение Эйлера – это система из n уравнений: $-\frac{d}{dt} f_{\dot{x}_i} + f_{x_i} = 0$, $i = 1, \dots, n$, каждое из которых является обыкновенным дифференциальным уравнением второго порядка. Порядок системы равен $2n$, следовательно, общее решение зависит от $2n$ произвольных постоянных, которые следует выбрать так, чтобы удовлетворить граничным условиям (9) (их число также равно $2n$).

Любое решение $x(t)$ уравнения Эйлера будем называть экстремалью задачи.

Пример 5. Найдем экстремали в задаче о брахистохроне:

$$f = \frac{\sqrt{1 + \dot{y}^2}}{\sqrt{y}}; \quad f_y = \frac{\sqrt{1 + \dot{y}^2}}{2\sqrt{y^3}}; \quad f_{\dot{y}} = \frac{\dot{y}}{\sqrt{1 + \dot{y}^2} \sqrt{y}}$$

$$\frac{d}{dx} f_{\dot{y}} = \frac{\ddot{y}}{\sqrt{(1 + \dot{y}^2)^3} \sqrt{y}} - \frac{\dot{y}^2}{2\sqrt{1 + \dot{y}^2} \sqrt{y^3}}$$

Поэтому уравнение Эйлера имеет вид

$$-\frac{\dot{y}}{\sqrt{y} \sqrt{(1 + \dot{y}^2)^3}} + \frac{\dot{y}^2}{2\sqrt{1 + \dot{y}^2} \sqrt{y^3}} - \frac{\sqrt{1 + \dot{y}^2}}{2\sqrt{y^3}} = 0.$$

или

$$2y\ddot{y} + \dot{y}^2 + 1 = 0.$$

Поскольку это уравнение не содержит x , его порядок понижается заменой $\dot{y} = p(y)$, $\ddot{y} = p \frac{dp}{dy}$:

$$2yp \frac{dp}{dy} + p^2 + 1 = 0; \quad \ln(1 + p^2) = \ln \frac{C}{y}; \quad p = \pm \sqrt{\frac{C}{y} - 1}$$

Взяв знак "+" получим $\frac{dy}{\sqrt{\frac{C}{y} - 1}} = dx$. Делая замену $y = C \sin^2 t$, имеем $x = \int \frac{2C \sin t \cos t dt}{\sqrt{\frac{1}{\sin^2 t} - 1}} = C \int 2 \sin^2 t dt$.

$$\begin{cases} x + D = Ct - C \frac{\sin 2t}{2} = \frac{C}{2} [2t - \sin 2t] \\ y = C \sin^2 t = \frac{C}{2} [1 - \cos 2t]. \end{cases} \quad (12)$$

Таким образом, экстремальными являются циклоиды.

Замечание 1. Строго говоря, предположения, при которых было выведено уравнение Эйлера, для задачи о брахистохроне не выполняются, т. к. при $y = 0$ подынтегральная функция терпит разрыв. Поэтому приведенные вычисления можно рассматривать как эвристический вывод. Строгое доказательство оптимальности решения (12) можно найти в [3], с. 24-27).

В заключение параграфа выпишем простейшие интегралы уравнений Эйлера.

1. Интеграл импульса. Если подынтегральная функция f не зависит явно от x_i , то $f_{\dot{x}_i}$ есть интеграл уравнений Эйлера. Действительно, i -е уравнение Эйлера имеет в этом случае вид

$$\frac{d}{dt} f_{\dot{x}_i} = 0$$

2. Интеграл энергии. Если подынтегральная функция f не зависит явно от t , то $H(x, \dot{x}) = \sum_{i=1}^n f_{\dot{x}_i} \dot{x}_i - f$ (или кратко $H = f_x \dot{x} - f$) есть интеграл уравнений Эйлера. Действительно,

$$H(x, \dot{x})|_{(11)} = \frac{d}{dt} (f_{\dot{x}}) \dot{x} + f_{\dot{x}} \ddot{x} - f_x \dot{x} - f_{\dot{x}} \ddot{x} = 0$$

(т. к., в силу (11), $\frac{d}{dt} f_{\dot{x}} = f_x$).

Названия этих интегралов взяты из классической механики. Они получают свое объяснение в следующей главе.

3 ГЕОДЕЗИЧЕСКИЕ НА РИМАНОВОМ МНОГООБРАЗИИ

Римановым многообразием называется многообразие M класса C^r , ($r > 1$), в каждом касательном пространстве которого определена структура евклидова пространства, т. е. задана положительно определенная квадратичная форма, определяющая скалярные произведения касательных векторов. При этом требуется, чтобы матрицы $g_{ij}(x)$ определяющие эту форму в некотором атласе многообразия M были функциями класса C^r . Длина кривой $x(t)$, $t \in [t_0, t_1]$, лежащей на многообразии, задается формулой:

$$\mathcal{L} = \int_{t_0}^{t_1} \sqrt{\sum_{i,j} g_{ij}(x(t)) \frac{dx_i}{dt} \frac{dx_j}{dt}} dt \quad (13)$$

(см. [5], с.33). Геодезическими линиями на M называются экстремали функционала (13). В том случае, когда многообразие вложено в евклидово пространство, на нем можно определить структуру риманова многообразия, индуцированную этим вложением. Например, для двумерной поверхности в \mathbb{R}^3 , задаваемой уравнением $r = r(u, v)$ или, что то же, $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$ метрика, индуцированная вложением, имеет вид $g_{11} = \langle r_u, r_u \rangle$, $g_{12} = g_{21} = \langle r_u, r_v \rangle$, $g_{22} = \langle r_v, r_v \rangle$. Выражение $ds^2 = \sum_{i,j} g_{ij} du_i du_j$ называется первой квадратичной формой поверхности ([6], с.218).

Пример 6 (геодезические на сфере). Без ограничения общности можно считать радиус сферы равным единице.

В сферических координатах $x = \cos \varphi \cos \psi$, $y = \sin \varphi \cos \psi$, $z = \sin \psi$ первая квадратичная форма имеет вид: $ds^2 = \cos^2 \psi d\varphi^2 + d\psi^2$. Примем ψ за независимую переменную, $\varphi = \varphi(\psi)$. Тогда

$$\mathcal{L} = \int \sqrt{\cos^2 \psi \cdot \dot{\varphi}^2 + 1} d\psi$$

Поскольку подынтегральная функция не зависит от φ , получаем интеграл импульса $\frac{\cos^2 \psi \cdot \dot{\varphi}}{\sqrt{\cos^2 \psi \cdot \dot{\varphi}^2 + 1}}$ или, что то же, $\dot{\varphi} = \frac{\pm C}{\sqrt{\cos^4 \psi - C^2 \cos^2 \psi}}$. Интегрируя, имеем $\varphi = \int \frac{\pm C dtg \psi}{\sqrt{1 - C^2 (tg^2 \psi + 1)}} = B \pm \arcsin \frac{C tg \psi}{\sqrt{1 - C^2}}$, т. е. $\sin(\varphi - B) = A tg \psi$. Отсюда: $A_1 \sin \varphi \cos \psi + A_2 \cos \varphi \cos \psi + A_3 \sin \varphi = 0$, и значит $A_1 x + A_2 y + A_3 z = 0$. Иными словами, геодезическими служат сечения сферы плоскостями, проходящими через начало координат, т. е. большие круги сферы.

Пример 7. (геодезические на плоскости Лобачевского). Модель Клейна-Пуанкаре плоскости Лобачевского (двумерного многообразия постоянной отрицательной кривизны) описывается следующим образом. Рассмотрим верхнюю полуплоскость $y > 0$ и введем на ней метрику, определяемую первой квадратичной формой

$$ds^2 = \frac{dx^2 + dy^2}{y^2}$$

Найдем геодезические. Принимая за независимое переменное x , получаем $\mathcal{L} = \int \frac{\sqrt{1 + \dot{y}^2}}{y} dx$. Поскольку подынтегральная функция не зависит явно от x , уравнение Эйлера имеет интеграл энергии $\frac{\dot{y}}{y\sqrt{1 + \dot{y}^2}} \dot{y} - \frac{\sqrt{1 + \dot{y}^2}}{y} = C$, т. е. $y\sqrt{1 + \dot{y}^2} = R$, или $dx = \frac{y dy}{\sqrt{R^2 - y^2}}$. Интегрируя, получаем $(x + D)^2 = R^2 - y^2$. Итак, геодезическими служат окружности, ортогональные прямой $y = 0$, которая в геометрии Лобачевского называется абсолютом. Если за независимое переменное принять y , то для подынтегральной функции $\frac{\sqrt{\dot{x}^2 + 1}}{y}$ можно записать интеграл импульса $\frac{\dot{x}}{y\sqrt{\dot{x}^2 + 1}} = C$, который, помимо прежних экстремалей, дает при $C = 0$ вертикальные прямые. Итак, роль "прямых" плоскости Лобачевского в модели Клейна-Пуанкаре играют прямые и окружности, ортогональные абсолюту (см. рис. 1).

Упражнение 1. Проверить, что для этих "прямых" выполняются все аксиомы евклидовой плоскости, кроме пятого постулата.

Упражнение 2. Показать, что гауссова кривизна многообразия с метрикой $ds^2 = \frac{dx^2 + dy^2}{y^2}$ постоянна и отрицательна.

Упражнение 3. Показать, что дробно-линейные преобразования, переводящие верхнюю полуплоскость в себя, сохраняют ds^2 и, следовательно, определяют группу движений плоскости Лобачевского.

Упражнение 4. Существует ли дробно-линейное преобразование, переводящее верхнюю полуплоскость в себя и переводящее две заданные точки этой полуплоскости в две другие заданные точки?

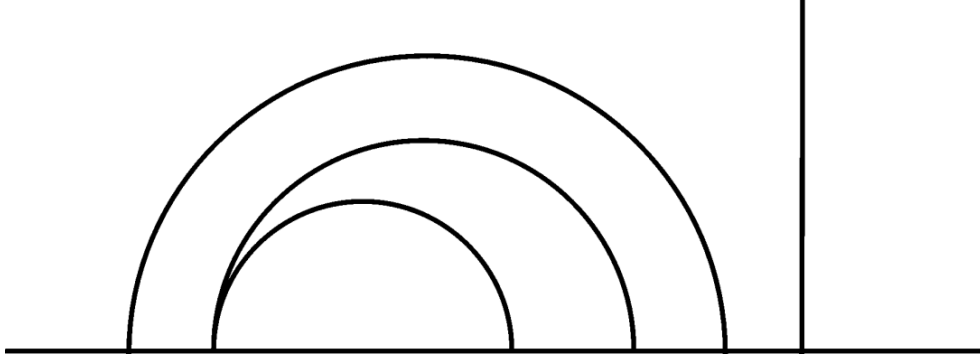


Рис. 1:

Вернемся к общей задаче нахождения экстремалей для функционала (13). Мы будем рассматривать кривые без особенностей: $|\dot{x}| \neq 0$.

Поскольку функционал (13) не зависит от выбора параметризации на кривой $x(t)$ мы можем считать, что за параметр выбирается длина дуги. В этом случае

$$\sum_{i,j} g_{ij}(x(t)) \dot{x}_i \dot{x}_j \equiv 1. \quad (14)$$

Найдем уравнение геодезических. С учетом (14) имеем

$$f_{\dot{x}_i} = \sum_{j=1}^n g_{lj} \dot{x}_j; \quad f_{x_l} = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial x_l} \dot{x}_i \dot{x}_j$$

Уравнение Эйлера имеет вид:

$$-\sum_{j=1}^n g_{lj} \ddot{x}_j - \frac{1}{2} \sum_{i,j=1}^n \frac{\partial g_{lj}}{\partial x_i} \dot{x}_i \dot{x}_j - \frac{1}{2} \sum_{i,j=1}^n \frac{\partial g_{il}}{\partial x_j} \dot{x}_i \dot{x}_j + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial x_l} \dot{x}_i \dot{x}_j = 0 \quad (15)$$

Обозначим через g^{kl} матрицу, обратную к матрице g_{ij} (g_{ij} положительно определена и, следовательно, не вырождена). Применим g^{kl} к обеим частям уравнения (15)

$$\ddot{x}_k + \frac{1}{2} \sum_{i,j,l=1}^n g^{kl} \left[\frac{\partial g_{lj}}{\partial x_i} + \frac{\partial g_{il}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_l} \right] \dot{x}_i \dot{x}_j = 0 \quad (16)$$

Выражение $\frac{1}{2} \sum_{l=1}^n g^{kl} \left[\frac{\partial g_{lj}}{\partial x_i} + \frac{\partial g_{il}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_l} \right]$ обыкновенно обозначают через Γ_{ij}^k и называют символом Кристоффеля или коэффициентами связности. В этих обозначениях уравнение (16) принимает вид:

$$\ddot{x}_k + \sum_{i,j=1}^n \Gamma_{ij}^k \dot{x}_i \dot{x}_j = 0, \quad k = 1, \dots, n. \quad (17)$$

Это и есть уравнение геодезических.

4 ВАРИАЦИЯ ФУНКЦИОНАЛА С ПОДВИЖНЫМИ КОНЦАМИ

Формула, которую мы сейчас выведем, будет играть определяющую роль во всем последующем изложении. Рассмотрим интегральный функционал

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

на однопараметрическом семействе кривых $x(t, \alpha)$ ($\alpha \in \mathbb{R}$); $x \in C^2(\mathbb{R} \times \mathbb{R})$, концы которых меняются с изменением α . Те величины, которые относятся к левому концу, мы будем помечать индексом 0, а те, которые относятся к правому – индексом 1. Пусть $t = t_0(\alpha)$, $x = x_0(\alpha)$ – гладкая кривая, по которой скользит левый конец; $t = t_1(\alpha)$, $x = x_1(\alpha)$ – гладкая кривая, по которой скользит правый конец.

Это означает, что

$$x(t_1(\alpha), \alpha) = x_1(\alpha); \quad x(t_0(\alpha), \alpha) = x_0(\alpha). \quad (18)$$

Имеем

$$\mathcal{J}(\alpha) = \int_{t_0(\alpha)}^{t_1(\alpha)} f(t, x(t, \alpha), \dot{x}(t, \alpha)) dt,$$

Вычислим $\frac{d\mathcal{J}}{d\alpha}(0)$, используя следующие обозначения: точка над переменной всегда обозначает дифференцирование по t .

$$x(t, 0) = \hat{x}(t), \quad \dot{x}(t, 0) = \hat{\dot{x}}, \quad t_1(0) = \hat{t}_1, \quad t_0(0) = \hat{t}_0,$$

$$\frac{\partial x}{\partial \alpha}(t_0) = h(t), \quad \frac{\partial \dot{x}}{\partial \alpha}(t, 0) = \dot{h}(t).$$

Последняя формула следует из теоремы о независимости смешанных производных от порядка дифференцирования. Тогда

$$\begin{aligned} \frac{d\mathcal{J}(\alpha)}{d\alpha} &= f(t_1(\alpha), x(t_1(\alpha), \alpha), \dot{x}(t_1(\alpha), \alpha)) \frac{dt_1}{d\alpha} - f(t_0(\alpha), x(t_0(\alpha), \alpha), \dot{x}(t_0(\alpha), \alpha)) \frac{dt_0}{d\alpha} + \\ &\quad \int_{t_0(\alpha)}^{t_1(\alpha)} (f_x \frac{\partial x}{\partial \alpha} + f_{\dot{x}} \frac{\partial \dot{x}}{\partial \alpha}) dt. \end{aligned} \quad (19)$$

Введем следующие сокращения в записи: $f(t, \hat{x}(t), \hat{\dot{x}}(t))$ будем обозначать через $\hat{f}(t)$ или просто \hat{f} . $f(t_0, \hat{x}(t_0), \hat{\dot{x}}(t_0))$ будем обозначать через \hat{f}_0 ; аналогично $\hat{f}_1 = f(t_1, \hat{x}(t_1), \hat{\dot{x}}(t_1))$. Символы \hat{f}_x , $\hat{f}_{\dot{x}}$, $(\hat{f}_{\dot{x}})_0$ и т. д. вводятся так же.

Подставив в (19) $\alpha = 0$, получим

$$\begin{aligned} \frac{d\mathcal{J}(0)}{d\alpha} &= \hat{f}_1 \frac{d\hat{t}_1}{d\alpha} - \\ &\quad \hat{f}_0 \frac{d\hat{t}_0}{d\alpha} + \int_{\hat{t}_0}^{\hat{t}_1} (\hat{f}_x h + \hat{f}_{\dot{x}} \dot{h}) dt. \end{aligned} \quad (20)$$

При выводе получаемой в этом параграфе формулы (23) будет использовано следующее

Предположение А. $\hat{f}_x(\cdot) \in C^1([t_0, t_1])$.

Проинтегрируем последнее слагаемое в (20) по частям:

$$\frac{d\mathcal{J}(0)}{d\alpha} = \hat{f}_1 \frac{dt_1}{d\alpha} - \hat{f}_0 \frac{dt_0}{d\alpha} + \hat{f}_x h \Big|_{\hat{t}_0}^{\hat{t}_1} + \int_{\hat{t}_0}^{\hat{t}_1} (\hat{f}_x - \frac{d}{dt} \hat{f}_x) h dt. \quad (21)$$

Преобразуем внеинтегральный член. Для этого продифференцируем тождества (18) по α при $\alpha = 0$:

$$\hat{x}(t_1) \frac{dt_1(0)}{d\alpha} + h(\hat{t}_1) = \frac{dx_1(0)}{d\alpha}; \quad \hat{x}(t_0) \frac{dt_0(0)}{d\alpha} + h(\hat{t}_0) = \frac{dx_0(0)}{d\alpha}.$$

Подставив $h(t_i)$, $i = 0, 1$ в (21), получим

$$\frac{d\mathcal{J}(0)}{d\alpha} = [\hat{f}_1 - (\hat{f}_x)_1 \hat{x}(t_1)] \frac{dt_1}{d\alpha} - [\hat{f}_0 - (\hat{f}_x)_0 \hat{x}(t_0)] \frac{dt_0}{d\alpha} + (\hat{f}_x)_1 \frac{dx_1(0)}{d\alpha} - (\hat{f}_x)_0 \frac{dx_0(0)}{d\alpha} + \int_{\hat{t}_0}^{\hat{t}_1} (\hat{f}_x - \frac{d}{dt} \hat{f}_x) h dt. \quad (22)$$

Формулу (22) удобно записать в следующей символической форме:

$$\begin{aligned} \frac{d\mathcal{J}(0)}{d\alpha} &= (pdx - Hdt) \Big|_{\hat{t}_0}^{\hat{t}_1} + \\ &\int_{\hat{t}_0}^{\hat{t}_1} (\hat{f}_x - \frac{d}{dt} \hat{f}_x) h dt. \end{aligned} \quad (23)$$

В этой записи подстановка верхнего предела \hat{t}_1 в выражение $pdx - Hdt$ означает не только подстановку \hat{t}_1 под знак канонических переменных p и H , но и замену (dt, dx) на касательный вектор $(\frac{dt_1}{d\alpha}, \frac{dx_1}{d\alpha})$ к кривой, по которой движется правый конец семейства; то же относится и к подстановке \hat{t}_0 .

Замечание 1. .

В дальнейшем формула (23) будет использоваться только в тех случаях, когда $\hat{x}(\cdot)$ – оптимальное решение, и, следовательно, предположение 1 будет выполнено в силу леммы Дюбуа Реймона. Поэтому (23) будет далее применяться без специальных оговорок.

Дифференциальная форма $pdx - Hdt$ называется формой Пуанкаре-Картана. Она играет фундаментальную роль в каноническом формализме и неоднократно будет использоваться в дальнейшем изложении.

5 УСЛОВИЯ ТРАНСВЕРСАЛЬНОСТИ

Рассмотрим задачу минимизации функционала \mathcal{J} с подвижными концами.

$$\Phi(t_0, x(t_0)) = 0, \quad \Psi(t_1, x(t_1)) = 0, \quad (24)$$

где $\Phi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^k$, $\Psi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^l$

Задача 1. . Среди всех кривых $x(\cdot)$ класса C^1 , удовлетворяющих условиям (24), найти кривую, которая дает функционалу \mathcal{J} локально минимальное значение.

Пусть $\hat{x}(t)$, $t \in [\hat{t}_0, \hat{t}_1]$ – решение этой задачи. Тогда та же кривая является решением задачи с закрепленными концами $\hat{x}(\hat{t}_0)$, $\hat{x}(\hat{t}_1)$ и, следовательно, для нее выполняется уравнение Эйлера. Однако, для нахождения \hat{t}_0 , \hat{t}_1 и еще $2n$ параметров, характеризующих искомую экстремаль (т. е. для нахождения $2n + 2$ параметров) мы имеем пока лишь $k + l$ граничных условий: $\Phi = 0$, $\Psi = 0$.

Для нахождения дополнительных граничных условий, которые должны выполняться на оптимальной траектории, рассмотрим произвольный касательный вектор $(\delta t_1, \delta x_1)$ к многообразию $\Psi = 0$ в точке $(\hat{t}_1, \hat{x}(\hat{t}_1))$. По определению касательного вектора на многообразии $\Psi = 0$ существует гладкая кривая $\gamma : \alpha \mapsto (t_1(\alpha), x_1(\alpha))$, касательная к которой совпадает с данным вектором, т. е. $(t_1(0), x_1(0)) = (\hat{t}_1, \hat{x}(\hat{t}_1))$; $(\frac{dt_1(0)}{d\alpha}, \frac{dx_1(0)}{d\alpha}) = (\delta t_1, \delta x_1)$. Обозначим через $x(t, \alpha)$ гладкое однопараметрическое семейство кривых, которое при $\alpha = 0$ содержит $\hat{x}(t)$; при $t = t_1(\alpha)$ дает кривую γ , а при $t = t_0(\alpha)$ дает точку $\hat{x}(\hat{t}_0)$, т. е. $x(t, 0) = \hat{x}(t)$, $x(t_1(\alpha), \alpha) = x_1(\alpha)$, $x(\hat{t}_0, \alpha) = \hat{x}(\hat{t}_0)$.

Возможность построения такого семейства (геометрически достаточно очевидная) формально обоснована в ([?], с.195). Подставляя это семейство под знак функционала и учитывая, что при $\alpha = 0$ \mathcal{J} достигает минимума, имеем

$$(pdx - Hdt)|_{\hat{t}_0}^{\hat{t}_1} + \int_{\hat{t}_0}^{\hat{t}_1} (f_x - \frac{d}{dt} f_{\dot{x}}) h dt = 0.$$

Поскольку для $\hat{x}(t)$ выполнено уравнение Эйлера, интегральный член равен нулю. Нулю равняется и результат подстановки \hat{t}_0 во внеинтегральный член, т. к. все кривые семейства проходят через точку $(\hat{t}_0, \hat{x}(\hat{t}_0))$ и, следовательно, $\frac{dt_0(0)}{d\alpha} = 0$, $\frac{dx_0(0)}{d\alpha} = 0$. Таким образом имеем, что

$$p_1 \delta x_1 - H_1 \delta t_1 = 0 \quad (25)$$

для любого вектора $(\delta t_1, \delta x_1)$, касательного к многообразию $\Psi = 0$ в точке $(\hat{t}_1, \hat{x}(\hat{t}_1))$. Совершенно аналогично получаем:

$$p_0 \delta x_0 - H_0 \delta t_0 = 0 \quad (26)$$

для любого вектора $(\delta t_0, \delta x_0)$ касательного к многообразию $\Phi = 0$ в точке $(\hat{t}_0, \hat{x}(\hat{t}_0))$.

Соотношения (25), (26) называются условиями трансверсальности.

Проверим, что соотношения (24), (25), (26) дают полную систему условий. Размерность многообразия $\Psi = 0$ (в регулярном случае, когда градиенты уравнений системы $\Psi = 0$ линейно независимы), равна $n + 1 - l$, и соотношения (25) дают $n + 1 - l$ независимых уравнений. Присоединив к ним l уравнений $\Psi = 0$, получаем на правом конце (так же, как и на левом) $n + 1$ уравнение. Это дает нужные $2n + 2$ соотношения.

Пример 1. Рассмотрим задачу минимизации функционала $\int_{t_0}^{t_1} \gamma(t, x) \sqrt{1 + \dot{x}^2} dt$, $x \in \mathbb{R}$, на кривых, ведущих из точки (t_0, x_0) на терминальное многообразие $x = \varphi(t)$. Пусть $(\delta t, \delta x)$ – касательный вектор к терминальному многообразию. Запишем соотношения (25) для нашей задачи: $p = \frac{\gamma \dot{x}}{\sqrt{1 + \dot{x}^2}}$; $H = \frac{-\gamma}{\sqrt{1 + \dot{x}^2}}$;

$$(p\delta x - H\delta t) = \frac{\gamma}{\sqrt{1 + \dot{x}^2}} (\dot{x}\delta x + \delta t) = 0 \quad (27)$$

Но $\frac{\delta x}{\delta t} = \varphi'(t)$ – тангенс угла наклона касательной. Поэтому условие (27): $\dot{x} \frac{\delta x}{\delta t} = -1$ есть условие ортогональности экстремали $x(\cdot)$ к терминальному многообразию $x = \varphi(t)$.

Пример 2. *Пример 11.* Рассмотрим задачу отыскания на римановом многообразии M кривой минимальной длины, соединяющей точку $x_0 \in M$ с подмногообразием $N \subset M$.

Длина кривой на M задается формулой

$$\mathcal{L} = \int_{t_0}^{t_1} \sqrt{\sum_{i,j=1}^n g_{ij}(x(t)) \frac{dx_i}{dt} \frac{dx_j}{dt}} dt,$$

где g_{ij} – метрический тензор. Выпишем канонические переменные для этого функционала:

$$p_j = \frac{\sum_{j=1}^n g_{ij}(x) \frac{dx_j}{dt}}{\sqrt{\sum_{i,j=1}^n g_{ij}(x) \frac{dx_i}{dt} \frac{dx_j}{dt}}},$$

$$H = -\sqrt{\sum_{i,j=1}^n g_{ij}(x) \frac{dx_i}{dt} \frac{dx_j}{dt}} + \frac{\sum_{i,j=1}^n g_{ij}(x) \frac{dx_i}{dt} \frac{dx_j}{dt}}{\sqrt{\sum_{i,j=1}^n g_{ij}(x) \frac{dx_i}{dt} \frac{dx_j}{dt}}} \equiv 0.$$

Условие трансверсальности (25) принимает вид:

$$\sum_{i=1}^n p_i \delta x_i, \text{ т. е. } \sum_{i,j=1}^n g_{ij}(x) \frac{dx_j}{dt} \delta x_i = 0.$$

Это означает, что вектор $\frac{dx}{dt}$ должен быть ортогонален (в смысле метрики g_{ij}) к любому вектору δx , касающемуся N в точке прихода экстремали $x(\cdot)$ на N . Итак, экстремаль должна подходить к N под прямым углом (в метрике g_{ij}).

6 ПРЕОБРАЗОВАНИЕ ЛЕЖАНДРА

Начнем с функций одного переменного. Рассмотрим класс \mathfrak{G} функций $f(u)$; $f : \mathbb{R} \rightarrow \mathbb{R}$, $f \in C^2(\mathbb{R})$, удовлетворяющих условиям:

1. $f''(u) > 0$, $u \in \mathbb{R}$.
2. Отображение $f' : \mathbb{R} \rightarrow \mathbb{R}$ сюръективно, т. е. его образ совпадает со всем пространством.

Каждой функции $f \in \mathfrak{G}$ поставим в соответствие новую функцию

$$f^*(p) = pu - f(u) \tag{28}$$

где новая независимая переменная p связана с u соотношением

$$p = \frac{df(u)}{du} \tag{29}$$

В силу условий 1, 2 соотношение (29) определяет взаимно однозначное соответствие между u и p . Именно поэтому функцию (28) можно считать функцией от p . Отображение $\Lambda : f \mapsto f^*$ называется преобразованием Лежандра.

а) Λ отображает \mathfrak{G} в \mathfrak{G} . Действительно,

$$\frac{df^*(p)}{dp} = u + p \frac{du}{dp} - \frac{df}{du} \frac{du}{dp}.$$

Два последних слагаемых взаимно уничтожаются в силу (29), поэтому

$$\frac{df^*}{dp} = u. \quad (30)$$

Дифференцируя (30) по p , получаем:

$$\frac{d^2 f^*}{dp^2} = \frac{du}{dp}.$$

Для нахождения $\frac{du}{dp}$ дифференцируем (29):

$$1 = \frac{d^2 f}{dp^2} \frac{du}{dp}; \quad \frac{du}{dp} = \left(\frac{d^2 f}{du^2} \right)^{-1}$$

(здесь использовано условие 1). Итак,

$$\frac{d^2 f^*}{dp^2} = \left(\frac{d^2 f}{du^2} \right)^{-1} > 0, \quad (31)$$

что доказывает выполнение для f^* условия 1. Выполнение условия 2 следует из формулы (30).

б) Отображение Λ инволютивно, т. е. $\Lambda^2 = I$, В самом деле, формула (30) определяет независимую переменную для функции $(f^*)^*$, при этом

$$(f^*)^*(u) = up - f^*(p) = up - [pu - f(u)] = f(u).$$

Геометрический смысл преобразования Лежандра можно пояснить следующим образом. Соотношение (29) означает, что мы ищем точку u , касательная в которой к графику функции $z = f(u)$ имеет наклон p . Формула (28) означает, что $f^*(p)$ есть величина, на которую надо опустить прямую $z = pu$, чтобы она стала касательной к графику функции $z = f(u)$. Тем самым, функция $f^*(p)$ определяет совокупность касательных к графику $z = f(u)$.

Обобщения

1. Все сказанное без труда переносится на случай функций многих переменных $f : \mathbb{R}^n \rightarrow \mathbb{R}$. При этом $p \in (\mathbb{R}^n)^*$; $f^*(p) = \langle p, u \rangle - f(u)$, $p_i = \frac{\partial f}{\partial u_i}$, $i = 1, \dots, n$.

Условия 1 и 2 принимают вид:

1. $f_{u_i u_j}$, т. е. гессиан функции f является положительно определенной матрицей.

7 КАНОНИЧЕСКИЕ ПЕРЕМЕННЫЕ

Как уже было сказано, система уравнений Эйлера

$$-\frac{d}{dt} f_{\dot{x}_i} + f_{x_i} = 0, \quad i = 1, \dots, n \quad (32)$$

имеет порядок $2n$. Поставим цель: записать эту систему в нормальной форме. Для этого естественно потребовать выполнение условия

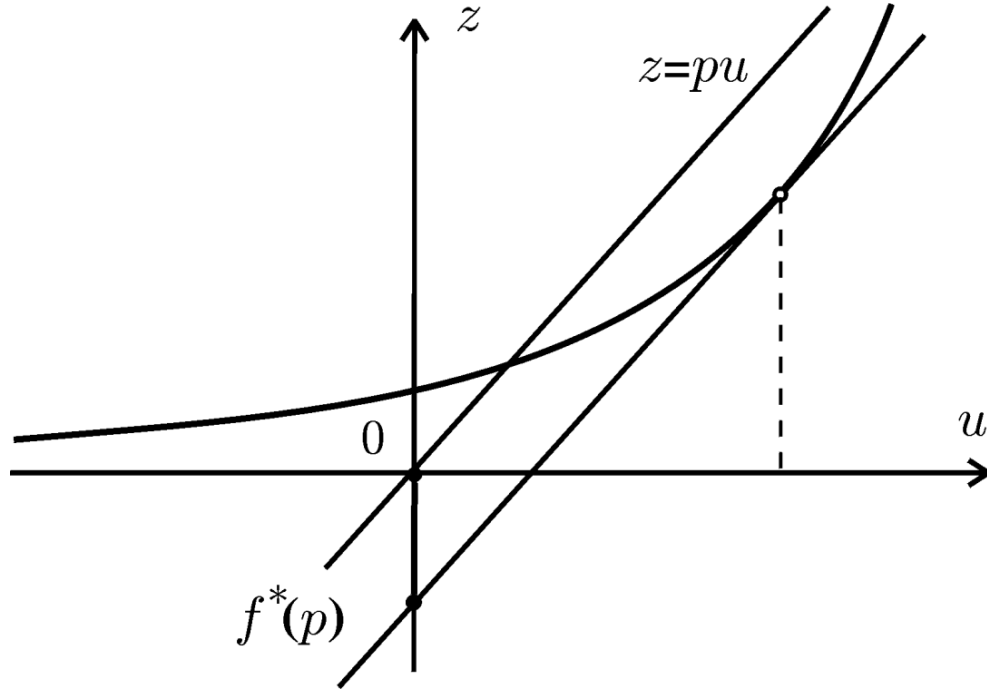


Рис. 2:

$$\det \left(\frac{\partial^2 f}{\partial \dot{x}_i \partial \dot{x}_j} \right) \neq 0, \quad (33)$$

которое называют условием Гильберта. Если оно выполнено, то систему (32) можно разрешить (локально) относительно вторых производные $\ddot{x}(t)$, а первые производные \dot{x} принять за новые неизвестные функции. Но есть и другой путь, предложенный английским математиком У. Р. Гамильтоном.

Предположим для простоты, что функция f , как функция от \dot{x} (при любых фиксированных значениях t, x), удовлетворяет условиям 1, 2 предыдущего параграфа. В частности, это означает, что условие Гильберта (33) заменяется на условие $\frac{\partial^2 f}{\partial \dot{x}_i \partial \dot{x}_j} > 0$. Преобразование Лежандра функции f (как функции от \dot{x} при фиксированных t и x) обозначим через $\mathbb{H}(t, x, p)$:

$$p_i = f_{\dot{x}_i}, \quad i = 1 \dots n; \quad \mathbb{H}(t, x, p) = \langle p, \dot{x} \rangle - f(t, x, \dot{x}).$$

Запишем в этих переменных систему (32). Найдем для этого

$$d\mathbb{H} = \langle dp, \dot{x} \rangle + \langle p, d\dot{x} \rangle - f_t dt - f_x dx - f_{\dot{x}} d\dot{x}.$$

Второе и последнее слагаемые взаимно уничтожаются, так что

$$d\mathbb{H} = \langle dp, \dot{x} \rangle - f_t dt - f_x dx$$

В силу теоремы об инвариантности первого дифференциала имеем

$$\frac{\partial \mathbb{H}}{\partial p} = \dot{x}, \quad \frac{\partial \mathbb{H}}{\partial t} = -f_t, \quad \frac{\partial \mathbb{H}}{\partial x} = -f_x. \quad (34)$$

Используя эти формулы, перепишем систему (32) в виде

$$\dot{x} = \mathbb{H}_p(t, x, p), \quad \dot{p} = -\mathbb{H}_x(t, x, p). \quad (35)$$

Системы вида (35) носят название канонических систем обыкновенных дифференциальных уравнений. Они обладают рядом замечательных свойств, с которыми можно познакомиться, например, в [8], с.57. Мы используем некоторые из этих свойств дальше в гл. VIII. В случае, когда f не удовлетворяет условиям 1, 2, канонические переменные вводятся с помощью преобразования Лежандра-Юнга.

8 МЕХАНИЧЕСКИЙ СМЫСЛ КАНОНИЧЕСКИХ ПЕРЕМЕННЫХ

Рассмотрим систему из N материальных точек $i = 1, \dots, N$ с массами m_i и с координатами $r_i = (x_i, y_i, z_i)$ в декартовой системе координат. Предположим, что потенциальная энергия этой системы задается функцией $U(r_1, \dots, r_N)$, кинетическая энергия $T = \frac{1}{2} \sum_{i=1}^n m_i \dot{r}_i^2$. В соответствии с принципом наименьшего действия движение системы происходит по экстремалиям функционала действия

$$\mathcal{L} = \int_{t_0}^{t_1} L(r, \dot{r}) dt, \quad (36)$$

где $L(z, \dot{r}) = T - U$. Экстремали – это решения уравнения Эйлера для функционала (36):

$$-\frac{d}{dt} L_{\dot{r}} + L_r = 0. \quad (37)$$

Система (37) состоит из $3N$ уравнений второго порядка. Запишем эти уравнения в канонической форме. Имеем $p_i = L_{\dot{r}_i} = m_i \dot{r}_i$, т. е. p_i в декартовой системе координат совпадает с импульсом материальной точки. Далее $\mathbb{H} = \sum_i \dot{r}_i (m_i \dot{r}_i) - T + U = T + U$ – это полная энергия системы (т. к.

L не зависит явно от t , то $\mathbb{H} = \text{const}$, т. е. полная энергия системы при движении сохраняется). Уравнения (37) в канонической форме принимают вид

$$\begin{aligned} \dot{r}_i &= \mathbb{H}_{p_i} = p_i / m_i \\ \dot{p}_i &= -\mathbb{H}_{r_i} = -\frac{\partial U}{\partial r_i}, \quad i = 1, \dots, N. \end{aligned}$$

Здесь первая группа уравнений соответствует определению импульса, а вторая – второму закону Ньютона, т. к. $-\frac{\partial U}{\partial r_i}$ – это сила, действующая на i -ю точку.

§

25. Условия Вейерштрасса-Эрдмана

Начнем с примера.

Пример 12. (Поверхности вращения с минимальной площадью). Найти кривую $y = y(x)$, соединяющую точки $(x_0, y_0) = a$ и $(x_1, y_1) = b$ и имеющую минимальную площадь поверхности вращения вокруг оси абсцисс.

Площадь поверхности вращения равна $2\pi \int y ds = 2\pi \int_{x_0}^{x_1} y \sqrt{1 + \dot{y}^2} dx$, поэтому нужно минимизировать функционал $\int y \sqrt{1 + \dot{y}^2} dx$ при условии $y(x_0) = y_0$, $y(x_1) = y_1$.

Интеграл энергии имеет вид $-y\sqrt{1+\dot{y}^2} + \frac{y\dot{y}^2}{\sqrt{1+\dot{y}^2}} = C$. Элементарные преобразования дают $y^2 = C^2 + C^2\dot{y}^2$ или $\int \frac{Cdy}{\sqrt{y^2-C^2}} = \int dx$.

После замены переменной $y = C \operatorname{ch} \tau$ получаем:

$$y = C \operatorname{ch} \frac{x+D}{C} \quad (38)$$

Если переписать минимизируемый функционал в виде

$$\int y ds = \int y(t) \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt,$$

то интеграл импульса по x : $\frac{y\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}} = C$, помимо уже найденных решений, при $C = 0$, дает вертикальные прямые $x = \operatorname{const}$.

Кривые (38) называются в механике "цепными линиями" поскольку именно такую форму принимает тяжелая гибкая однородная нерастяжимая нить, подвешенная за концы. Для того, чтобы пояснить этот факт, достаточно заметить, что формула (38) определяет координату центра тяжести однородной кривой, а нить стремится занять такое положение, при котором эта координата минимальна. Правда, надо минимизировать этот функционал на множестве кривых заданной длины (нить нерастяжима). Полное решение этой задачи требует применения правила множителей Лагранжа, которое следует изучить по книге [7].

Поверхность вращения цепной линии носит название катеноид. Посмотрим, всегда ли можно провести кривую (38), соединяющую точки a и b . Пусть, например, $a = (q, 1)$, $b = (-q, 1)$. Тогда (вследствие симметрии граничных условий относительно начала координат) $D = 0$, а C надо найти как решение уравнения $C \operatorname{ch} \frac{q}{C} = 1$. Обозначим $1/C = z$ и будем решать уравнение $\operatorname{ch} qz = z$.

На рис. 3 изображены графики кривых $w = z$ и $w = \operatorname{ch} qz$ для различных значений q . При достаточно большом q , эти кривые не пересекаются, т. е. в этом случае не существует решения уравнений Эйлера, проходящего через точки a и b . Этот факт требует осознания.

Пример 12 является частным случаем задачи Плато, которая ставится следующим образом: среди всех поверхностей, имеющих в качестве своей границы данный контур K , найти такую, которая имеет минимальную площадь. С естественно-научной точки зрения речь идет о том, чтобы найти форму мыльной пленки, затягивающей контур K (мыльная пленка, вследствие сильного поверхностного натяжения, стремится занять положение с минимальной площадью поверхности). Пример 12 соответствует случаю, когда контур K состоит из двух окружностей, полученных при вращении точек a и b вокруг оси абсцисс. Решением, как мы видели, служит катеноид. Но как объяснить отсутствие решения при больших значениях q ?

Для того, чтобы ответить на этот вопрос, обратимся к общей теории. Расширим класс допустимых кривых в задаче 5 и будем искать минимум функционала не в классе гладких кривых, а в классе кусочно-гладких кривых $\mathcal{KC}^1([t_0, t_1])$

Определение. Пространство $\mathcal{KC}^1([t_0, t_1])$ – это множество непрерывных на $[t_0, t_1]$ функций $x(t)$, для каждой из которых существует конечное число точек $\tau_1, \dots, \tau_n \in [t_0, t_1]$ таких, что ограничение $x(t)$ на любой из отрезков $[t_0, \tau_1], \dots, [\tau_i, \tau_{i+1}], \dots, [\tau_n, t_1]$ является гладкой функцией на этом отрезке. Это множество снабжается топологией подпространства пространства $C[t_0, t_1]$.

Очевидно, что для решения $\hat{x}(\cdot) \in \mathcal{KC}^1([t_0, t_1])$ задачи 5, каждый участок гладкости $\hat{x}(t)$, $t \in [\tau_i, \tau_{i+1}]$ должен удовлетворять уравнению Эйлера, поскольку он должен давать минимум функционалу \mathcal{J} с граничными условиями $(\tau_i, \hat{x}(\tau_i))$, $(\tau_{i+1}, \hat{x}(\tau_{i+1}))$. Однако, большинство таких

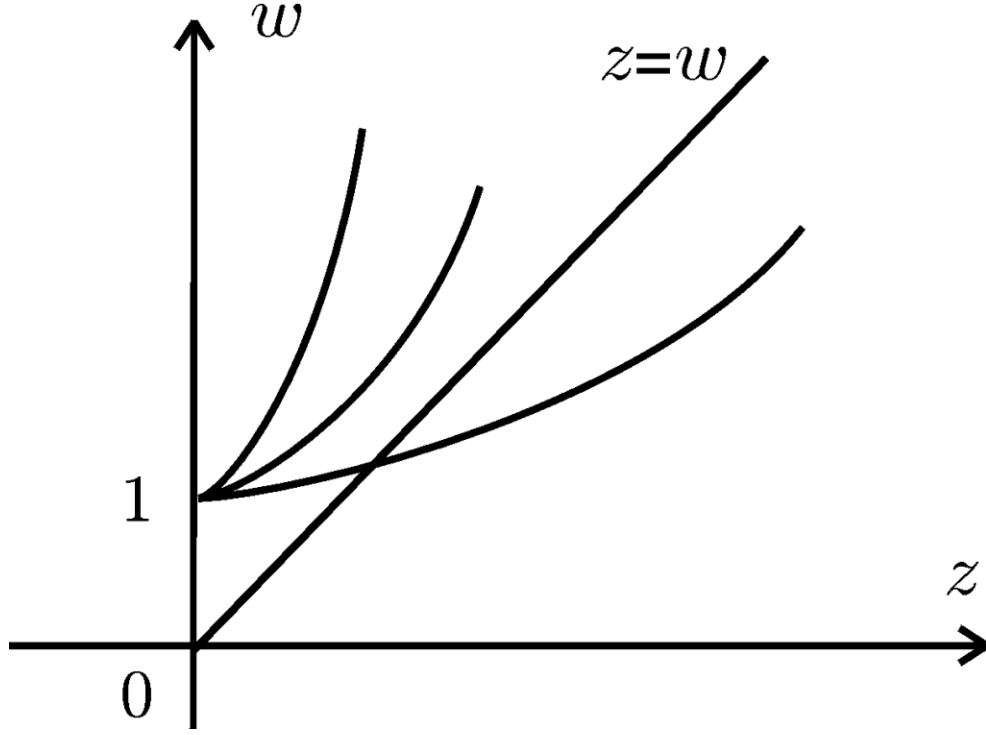


Рис. 3:

экстремалей заведомо не дает минимума функционалу \mathcal{J} . Оказывается, что в каждой угловой точке должно выполняться следующее дополнительное условие.

Условие Вейерштрасса-Эрдмана. Канонические переменные в угловой точке должны быть непрерывными.

Нетривиальность этого условия состоит в том, что в выражение, определяющее канонические переменные, входит \dot{x} , которое в угловой точке терпит разрыв.

Доказательство условия Вейерштрасса-Эрдмана. Очевидно, что достаточно рассмотреть случай одной угловой точки. Пусть $\hat{x}(\cdot)$ имеет точку излома $(\tau, \hat{x}(\tau))$. Проведем через эту точку произвольную прямую l с направляющим вектором $(\delta t, \delta x)$. Так же, как в § 24, рассмотрим гладкое однопараметрическое семейство кривых $x(t, \alpha)$ такое, что $x(t, 0) = \hat{x}(t)$; $x(t_0, \alpha) = \hat{x}(t_0)$; $x(t_1, \alpha) = \hat{x}(t_1)$ и при фиксированном α функция $x(t, \alpha)$ является гладкой всюду, кроме точки пересечения с прямой l .

Применим формулу (?), разбив $\mathcal{J}(\alpha)$ на два слагаемых, до и после пересечения с l :

$$\begin{aligned} \mathcal{J}(\alpha) &= \int_{t_0}^{\tau(\alpha)} f(t, x(t, \alpha), \dot{x}(t, \alpha)) dt + \int_{\tau(\alpha)}^{t_1} f(t, x(t, \alpha), \dot{x}(t, \alpha)) dt, \\ \frac{d\mathcal{J}(0)}{d\alpha} &= [p(\tau - 0)\delta x - \mathbb{H}(\tau - 0)\delta t] - [p(\tau + 0)\delta x - \mathbb{H}(\tau + 0)\delta t] = \\ &= [p(\tau - 0) - p(\tau + 0)]\delta x - [\mathbb{H}(\tau - 0) - \mathbb{H}(\tau + 0)]\delta t = 0. \end{aligned} \quad (39)$$

В этой формуле интегральный член отсутствует в силу выполнения уравнения Эйлера на $\hat{x}(t) = x(t, 0)$, а подстановка t_0 и t_1 также дает 0, т. к. все кривые $x(t, \alpha)$ проходят через точки $(t_0, \hat{x}(t_0))$ и $(t_1, \hat{x}(t_1))$. Поскольку δt и δx произвольны, из (39) следует непрерывность канонических переменных в точке τ .

Вернемся теперь к примеру 12 и попробуем найти кусочно-гладкие решения задачи о мини-

мальной поверхности вращения. Выпишем для этого канонические переменные:

$$p = \frac{y\dot{y}}{\sqrt{1 + \dot{y}^2}}; \quad \mathbb{H} = \frac{-y}{\sqrt{1 + \dot{y}^2}}$$

Нетрудно показать, что p и \mathbb{H} могут остаться непрерывными при скачке \dot{y} тогда и только тогда, когда $y = 0$. Но ни одна из линий (38) не доходит до оси абсцисс. Поэтому ломанными экстремалиями могут служить только линии, состоящие из отрезка оси абсцисс и двух вертикальных отрезков $x = \text{const}$. Следовательно, ответ на вопрос, что случится с мыльной пленкой, если две окружности, составляющие контур K , отодвигать друг от друга, увеличивая расстояние q между ними, очень прост: пленка лопнет и затянёт две окружности. Это и есть поверхность вращения найденной ломаной экстремали. Отрезок оси вращения, соединяющий два диска имеет при этом нулевую площадь и служит лишь напоминанием о связности катеноида. Оказывается, что в общей задаче Плато такого рода клетки меньшей размерности играют весьма важную роль (см. об этом в [10]).

§

26. Уравнение Гамильтона-Якоби

Рассмотрим задачу 5 с несколько иной точки зрения. Зафиксируем начальную точку $(t_0, x_0) = a$ и введем функцию $S(t_1, x_1)$, равную минимальному значению функционала \mathcal{J} на траекториях, соединяющих точки $a = (t_0, x_0)$ и $b = (t_1, x_1)$. Найдем уравнение в частных производных, которому удовлетворяет функция S .

Рассмотрим произвольное гладкое однопараметрическое семейство траекторий $x(t, \alpha)$ с неподвижным левым и подвижным правым концом, такое, что $x(t, 0) = \hat{x}(t)$. Применяя к этому семейству формулу (38) получим $\frac{\partial S}{\partial \alpha} = p \frac{dx_1}{d\alpha} - \mathbb{H} \frac{dt_1}{d\alpha}$. Учитывая произвольность кривых, по которым скользит правый конец, получим $dS = p dx_1 - \mathbb{H} dt_1$.

Следовательно, дифференциал функции $S(t, x)$ равен форме Пуанкаре-Картана.

Начиная с этого момента будем опускать индекс 1 у переменных (t, x) . Имеем $\frac{\partial S}{\partial x}(t, x) = p$, $\frac{\partial S}{\partial t}(t, x) = -\mathbb{H}(t, x, p)$. Отсюда следует, что

$$\frac{\partial S}{\partial t}(t, x) + \mathbb{H}(t, x, \frac{\partial S}{\partial x}(t, x)) = 0. \quad (40)$$

Это и есть уравнение Гамильтона-Якоби.

Пример 13. Запишем уравнение Гамильтона-Якоби для функционала

$$\int_{t_0}^{t_1} \gamma(t, x) \sqrt{1 + \sum_{i=1}^n \dot{x}_i^2} dt. \quad (41)$$

Гамильтониан для функционала (41) был найден в примере 9:

$$\mathbb{H}(t, x, p) = -\sqrt{\gamma^2 - \sum_{i=1}^n p_i^2}.$$

Уравнение Гамильтона-Якоби имеет вид

$$\frac{\partial S}{\partial t} = \sqrt{\gamma^2 - \sum_{i=1}^n \left(\frac{\partial S}{\partial x_i}\right)^2} \quad \text{или} \quad \left(\frac{\partial S}{\partial t}\right)^2 + \sum_{i=1}^n \left(\frac{\partial S}{\partial x_i}\right)^2 = \gamma^2(t, x).$$

Это уравнение называется уравнением Эйконала для неоднородной среды. (Ср. пример 2, с.16).

Упражнение. Найти уравнение Гамильтона-Якоби для функционала действия на римановом многообразии

$$\int \sum_{i,j=1}^n g_{ij}(x) \frac{dx_i}{dt} \frac{dx_j}{dt} dt.$$

9 УСЛОВИЕ ЛЕЖАНДРА

Определение 1. Будем говорить, что $\delta^2 \mathcal{J}$ удовлетворяет условию Лежандра, если для всех $t \in [t_0, t_1]$ матрица $A(t)$ неотрицательно определена ($A(t) \geq 0$).

Теорема 1. Для неотрицательной определенности $\delta^2 \mathcal{J}$ необходимо выполнение условия Лежандра.

Доказательство: Пусть $\tau \in (t_0, t_1)$ и в точке τ нарушается условие Лежандра, т. е. существует такой вектор $\xi \in \mathbb{R}^n$, что $\langle A\xi, \xi \rangle < 0$. Возьмем $\sigma > 0$ столь малым, чтобы $t_0 < \tau - \sigma < \tau + \sigma < t_1$, и рассмотрим функцию $\chi(t)$ (см. рис. 4)

$$\chi(t) = \begin{cases} \sigma - |t - \tau| & \text{при } |t - \tau| < \sigma \\ 0 & \text{при } |t - \tau| \geq \sigma \end{cases}$$

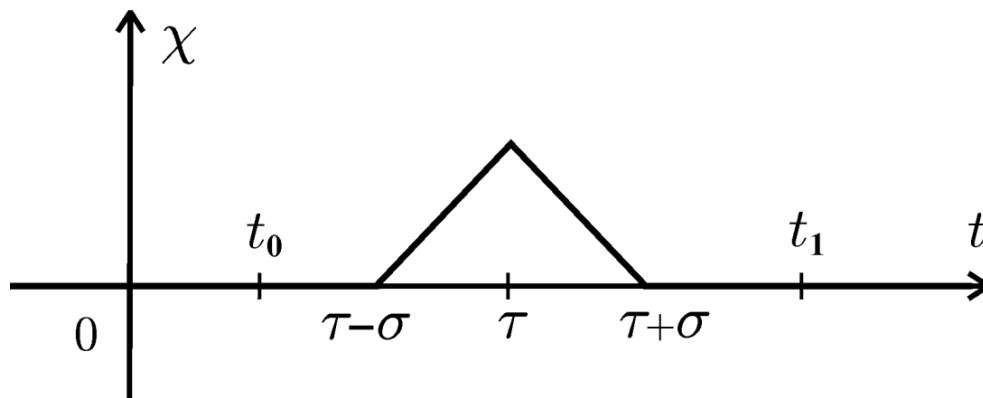


Рис. 4:

Вычислим значение $\delta^2 \mathcal{J}$ на функции $h(t) = \chi(t)\xi$: $\delta^2 \mathcal{J}(\chi(t)\xi) = \int_{\tau-\sigma}^{\tau+\sigma} (\langle A\xi, \xi \rangle + 2 \operatorname{sgn}(\tau - t) \langle C\xi, \xi \rangle \chi(t) + \langle B\xi, \xi \rangle \chi^2(t)) dt$.

Рассмотрим сначала первое слагаемое. Поскольку $A(t)$ непрерывна и $\langle A(\tau)\xi, \xi \rangle < 0$ существует такое $k > 0$, что $\langle A(t)\xi, \xi \rangle < -k$ при $t \in [\tau - \sigma, \tau + \sigma]$ для достаточно малых σ . Все остальные слагаемые в подынтегральном выражении равномерно стремятся к нулю при $\sigma \rightarrow 0$, следовательно $\delta^2 \mathcal{J} < 0$ при достаточно малом σ .

Доказательство на этом еще не завершается, поскольку функционал $\delta^2 \mathcal{J}$ оказался отрицательным не на гладкой, а на кусочно-гладкой функции. Для завершения доказательства следует воспользоваться леммой о округлении углов ([?] с.69) в которой доказывается, что значение интегрального функционала на любой кусочно-гладкой функции можно сколь угодно точно приблизить его значениями на гладких функциях с теми же граничными значениями. Так что если $\delta^2 \mathcal{J}$ на некоторой кусочно-гладкой функции принимает отрицательное значение, то найдется гладкая

функция, на которой он отрицателен. А это уже противоречит неотрицательной определенности $\delta^2 \mathcal{J}$.

Мы доказали неотрицательность матрицы $A(t)$ во внутренних точках отрезка $[t_0, t_1]$. Ее неотрицательность на концах является следствием ее непрерывности.

Возвращаясь к задаче 5, получаем в качестве следствия необходимое условие локального минимума функционала

Теорема 1. . (Необходимое условие Лежандра). Если экстремаль $\hat{x}(\cdot)$ дает минимум функционалу \mathcal{J} в задаче 5, то при всех $t \in [t_0, t_1]$ матрица $\hat{f}_{\dot{x}_i \dot{x}_j}$ неотрицательно определена.

Определение 1. .

Будем говорить, что $\delta^2 \mathcal{J}$ удовлетворяет усиленному условию Лежандра, если для всех $t \in [t_0, t_1]$ матрица $A(t)$ положительно определена ($A(t) > 0$).

Все дальнейшие исследования второй вариации $\delta^2 \mathcal{J}$ будем по традиции, восходящей к Якоби, проводить только в случае, когда выполнено усиленное условие Лежандра. -

10 СОПРЯЖЕННЫЕ ТОЧКИ

Для изучения вопроса о неотрицательной определенности квадратичного функционала $\delta^2 \mathcal{J}$ воспользуемся вышеизложенными результатами вариационного исчисления. Заметим, что неотрицательная определенность функционала $\delta^2 \mathcal{J}$ эквивалентна тому, что $h(\cdot)$ дает минимум функционалу $\delta^2 \mathcal{J}$. Поэтому рассмотрим следующую, так называемую присоединенную, экстремальную задачу: $\delta^2 \mathcal{J} = \int_{t_0}^{t_1} [\langle A\dot{h}, \dot{h} \rangle + 2\langle C\dot{h}, h \rangle + \langle Bh, h \rangle] dt \rightarrow \inf$; $h(\cdot) \in \mathbb{C}$. Запишем уравнение Эйлера для присоединенной задачи

$$-\frac{d}{dt}[A\dot{h} + C^*h] + [C\dot{h} + Bh] = 0. \quad (42)$$

Уравнение (42) называется уравнением Якоби для исходной задачи.

Упражнение. Показать, что уравнение Якоби – это уравнение в вариациях для уравнения Эйлера.

Лемма 1. Функционал $\delta^2 \mathcal{J}$ обращается в ноль на любом решении $h(\cdot) \in \mathbb{C}$ уравнения (42).

Доказательство. Запишем $\delta^2 \mathcal{J}$ в виде: $\delta^2 \mathcal{J} = \int_{t_0}^{t_1} [\langle (A\dot{h} + C^*h), \dot{h} \rangle + \langle (C\dot{h} + Bh), h \rangle] dt$. Интегрируя первый член по частям и учитывая, что внеинтегральный член обращается в ноль, поскольку $h(\cdot) \in \mathbb{C}([t_0, t_1])$, получаем

$$\int_{t_0}^{t_1} \langle [-\frac{d}{dt}(A\dot{h} + C^*h) + (C\dot{h} + Bh)], h \rangle dt = 0,$$

причем последнее равенство следует из того, что $h(\cdot)$ – решение уравнения Якоби.

Рассмотрим матричное решение $U(t)$ уравнения (42):

$$-\frac{d}{dt}[AU' + C^*U] + [CU' + BU] = 0. \quad (43)$$

где $U(t)$ – $(n \times n)$ -матрица, удовлетворяющая начальным условиям:

$$U(t_0) = 0, \quad U'(t_0) = 1. \quad (44)$$

Это означает, что i -й столбец матрицы $U(t)$ есть решение $h_i(t)$ уравнения (42), удовлетворяющее условиям $h_i(t_0) = 0$; $\dot{h}_i(t_0) = e_i$, где e_i – единичный базисный орт.

Определение 2. .

Точка $\tau > t_0$ называется сопряженной с точкой t_0 , если существует нетривиальное решение $\tilde{h}(t)$ уравнения (42) такое, что $\tilde{h}(t_0) = \tilde{h}(\tau) = 0$.

Предложение 1. .

Точка $\tau > t_0$ является сопряженной с точкой t_0 тогда и только тогда, когда $\det U(\tau) = 0$.

Доказательство.

Пусть $\det U(\tau) = 0$. Тогда существует нетривиальный набор констант c_i , $i = 1, \dots, n$, такой, что линейная комбинация столбцов $h_i(\tau)$ матрицы $U(\tau)$ с коэффициентами c_i равна нулю: $\sum c_i h_i(\tau) = 0$. Рассмотрим функцию $\tilde{h}(t) = \sum c_i h_i(t)$, которая является решением системы (42). Это решение не тривиально, т. к. в точке t_0 его производная отлична от нуля, и оно обращается в ноль как при $t = t_0$, так и при $t = \tau$.

Пусть теперь точка τ является сопряженной в смысле определения 3. Тогда решение $\tilde{h}(\cdot)$ можно выразить как линейную комбинацию решений $h_i(\cdot)$, разложив $\tilde{h}_i(t_0)$ по базису e_i : $\tilde{h}(t) = \sum c_i h_i(t)$, при этом набор c_i не тривиален. Поскольку $\tilde{h}(\tau) = 0$, мы имеем нетривиальную линейную комбинацию столбцов матрицы $U(\tau)$, дающую нулевой столбец. Следовательно, $\det U(\tau) = 0$.

□

11 НЕОТРИЦАТЕЛЬНАЯ ОПРЕДЕЛЕННОСТЬ ВТОРОЙ ВАРИАЦИИ

Теорема 2. .

Для неотрицательной определенности квадратичного функционала $\delta^2 \mathcal{J}$ при условии выполнения усиленного условия Лежандра ($A(t) > 0$) необходимо, чтобы интервал (t_0, t_1) не содержал точек, сопряженных с точкой t_0 .

Доказательство.

Предположим противное. Пусть $\tau \in (t_0, t_1)$ – точка, сопряженная с точкой t_0 . Тогда в силу определения 3 существует нетривиальное решение $\tilde{h}(t)$ уравнения Якоби (18), обращающееся в ноль при $t = t_0$ и при $t = \tau$: $\tilde{h}(t_0) = \tilde{h}(\tau) = 0$. Продолжим эту функцию нулем на отрезок $[\tau, t_1]$:

$$\bar{h}(t) = \begin{cases} \tilde{h}(t) & \text{при } t \in [t_0, \tau] \\ 0 & \text{при } t \in [\tau, t_1] \end{cases}$$

Применив лемму 1 к функции $\bar{h}(t)$ и отрезку $[t_0, \tau]$, нетрудно убедиться, что $\delta^2 \mathcal{J}(\bar{h}(\cdot)) = 0$. Поскольку $\delta^2 \mathcal{J} \geq 0$, получаем, что $\bar{h}(\cdot)$ есть решение присоединенной задачи, имеющее точку излома τ . Следовательно, в этой точке должно выполняться условие Вейерштрасса-Эрдмана. Запишем это условие. Импульс в присоединенной задаче равен $2A\dot{h} + 2C^*h$. Поскольку $\tilde{h}(\tau) = 0$, а матрица A невырождена, непрерывность импульса влечет за собой непрерывность \dot{h} . Но $\dot{h}(\tau +$

$0) = 0$. Следовательно, $\dot{\tilde{h}}(\tau-0) = 0$. Функция $\tilde{h}(\cdot)$ на отрезке $[t_0, \tau]$ является решением уравнения второго порядка (18) и в точке τ удовлетворяет условиям $\tilde{h}(\tau) = 0$, $\dot{\tilde{h}}(\tau) = 0$. Следовательно, $\tilde{h}(\tau) \equiv 0$, что противоречит определению сопряженной точки. Теорема доказана.

□

Непосредственным следствием теоремы является

Теорема 3 (Необходимое условие Якоби). .

Пусть $\hat{x}(\cdot)$ – решение задачи 5. Тогда, если на $\hat{x}(\cdot)$ выполнено усиленное условие Лежандра ($\hat{f}_{\dot{x}\dot{x}} > 0$), то интервал (t_0, t_1) не содержит точек, сопряженных с t_0 .

Определение 3. .

Будем говорить, что для функционала $\delta^2 \mathcal{J}$ выполнено условие Якоби, если интервал (t_0, t_1) не содержит точек, сопряженных с t_0 . Будем говорить, что выполнено усиленное условие Якоби, если полуинтервал $(t_0, t_1]$ не содержит точек, сопряженных с t_0 .

12 УРАВНЕНИЕ РИККАТИ

Для начала рассмотрим простейшую задачу классического вариационного исчисления. Минимизировать функционал

$$\int_{t_0}^{t_1} f(t, x(t), \dot{x}(t)) dt, \quad x(t_0) = a, x(t_1) = b \quad (45)$$

на пространстве гладких функций $x(t)$. Будем считать, что подынтегральная функция $f : [t_0, t_1] \times \mathbb{R}^n \times \mathbb{R}^n$ и искомая функция $x(\cdot) : [t_0, t_1] \rightarrow \mathbb{R}^n$ гладкие. Первая вариация, т.е. главная линейная часть приращения функционала (88), вычисленная на траектории, дающей C^1 -минимум функционалу (88), должна равняться нулю. Это приводит к уравнению Эйлера

$$-\frac{d}{dt}(f_{\dot{x}}) + f_x = 0. \quad (46)$$

Решения уравнения Эйлера называются экстремалиями. Вторая вариация, т.е. главная квадратичная часть отклонения функционала (88) от касательной плоскости, вычисленная на экстремали $\hat{x}(\cdot)$, имеет вид

$$\frac{1}{2} \int_{t_0}^{t_1} \{ \langle \hat{f}_{\dot{x}\dot{x}} \dot{h}, \dot{h} \rangle + 2 \langle \hat{f}_{\dot{x}x} \dot{h}, h \rangle + \langle \hat{f}_{xx} h, h \rangle \} dt, \quad h(t_0) = h(t_1) = 0. \quad (47)$$

Здесь и в дальнейшем значок $\hat{\cdot}$ над знаком функции означает, что в нее подставлена экстремаль $\hat{x}(\cdot)$. Матрицы квадратичных форм $A = \hat{f}_{\dot{x}\dot{x}}$ и $B = \hat{f}_{xx}$ симметрические ($A^* = A, B^* = B$), матрица билинейной формы C произвольная.

Необходимое условие оптимальности экстремали $\hat{x}(\cdot)$ состоит в том, чтобы вторая вариация (90) была неотрицательно определена на пространстве $C^1[t_0, t_1]$. Для этого в свою очередь необходимо, чтобы матрица $A(t)$ была неотрицательно определена при всех $t \in [t_0, t_1]$ (условие Лежандра). Положительная определенность второй вариации есть достаточное условие слабого минимума, т.е. минимума в смысле метрики C^1 .

Будем предполагать, что выполнено усиленное условие Лежандра, т.е. что матрица $A(t)$ положительно определена. Для того, чтобы сделать все подынтегральное выражение положительным,

вычтем под интегралом (90) производную от квадратичной формы $\langle W(t)h, h \rangle$. Значение функционала не изменится, т.к. на концах отрезка функция h обращается в ноль. Подберем матрицу W так, чтобы под интегралом осталась только квадратичная форма с матрицей A . Под интегралом получим

$$\begin{aligned} & \langle A\dot{h}, \dot{h} \rangle + 2 \langle C\dot{h}, h \rangle + \langle Bh, h \rangle - 2 \langle W\dot{h}, h \rangle - \langle \dot{W}h, h \rangle = \\ & \quad \left(A^{1/2}\dot{h} + A^{-1/2}(C^* - W)h \right) - \\ & \quad \langle \dot{W}h, h \rangle - \langle A^{-1}(C^* - W)h, (C^* - W)h \rangle + \langle Bh, h \rangle. \end{aligned}$$

Получаем уравнение

$$\dot{W} + (C - W)A^{-1}(C^* - W) - B = 0, \quad (48)$$

которое называется матричным уравнением Риккати. Итак, для того чтобы получить под интегралом положительную функцию, надо найти решение уравнения (91), определенное на всем отрезке $[t_0, t_1]$. Это не всегда возможно, потому что правая часть уравнения квадратична и решение может уйти в бесконечность.

13 ГРАССМАНОВЫ МНОГООБРАЗИЯ

Рассмотрим совокупность линейных подпространств линейного пространства \mathbb{R}^n (точно так же можно рассмотреть случаи комплексного пространства \mathbb{C}^n , или пространства кватернионов \mathbb{H}^n). Покажем, что эта совокупность образует гладкое многообразие. Карты этого многообразия будем задавать парами плоскостей, находящихся в общем положении: k -мерной плоскости V_0 , называемой горизонтальной, и $(n - k)$ -мерной плоскости V_∞ , называемой вертикальной ($V_0 \cap V_\infty = 0$). Будем считать, что подпространство W принадлежит данной карте, если $W \cap V_\infty = 0$. Рассмотрим два проектора $\text{pr}_0 : W \rightarrow V_0$ и $\text{pr}_\infty : W \rightarrow V_\infty$ параллельно плоскостям V_∞ и V_0 соответственно. По условию, оператор pr_0 обратим, т.к. размерности V_0 и W одинаковы и $W \cap V_\infty = 0$. За координаты подпространства W примем матрицу отображения

$$W = \text{pr}_\infty \circ \text{pr}_0^{-1} : V_0 \rightarrow V_\infty. \quad (49)$$

Саму плоскость и ее координату, являющуюся $(k \times (n - k))$ -матрицей, будем обозначать одной и той же буквой. Очевидно, что любая $(k \times (n - k))$ -матрица W определяет линейное подпространство W как множество точек (x, Wx) , $x \in V_0$. Найдем функции перехода. Пусть W лежит в пересечении карт V_0, V_∞ и $\tilde{V}_0, \tilde{V}_\infty$. Запишем преобразование \mathbb{R}^n , которое переводит V_0 в \tilde{V}_0 , а V_∞ в \tilde{V}_∞ в блочной форме

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

где A есть $(k \times k)$ -матрица. Множество точек (x, Wx) перейдет в

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ Wx \end{pmatrix} = \begin{pmatrix} (Ax + BWx) \\ (Cx + DWx) \end{pmatrix}$$

Обозначим $(A + BW)x := y$. Матрица $A + BW$ не вырождена, т.к. иначе образ вектора $x \in W$ имел бы ненулевую вертикальную составляющую в карте $\tilde{V}_0, \tilde{V}_\infty$. Поэтому можно считать, что $x = (A + BW)^{-1}y$ и образ подпространства W есть $\tilde{W} = (C + DW)(A + BW)^{-1}$. Таким образом

переход от карты к карте задается матричным дробно линейным преобразованием, которое называется матричным преобразованием Мебиуса. Тем самым, множество k -мерных подпространств n -мерного линейного пространства снабжается структурой гладкого (и даже алгебраического) многообразия. Полученное многообразие обозначается $\text{Gr}(k, \mathbb{R}^n)$ (или $\text{Gr}(k, n)$, если подразумевается, что основное поле есть \mathbb{R}) и называется многообразием Грассмана.

Композиция преобразований Мебиуса с матрицами

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

и

$$\begin{pmatrix} K & L \\ M & N \end{pmatrix}$$

приводит к преобразованию

$$\{M + N[(C + DW)(A + BW)^{-1}]\} \{K + L[(C + DW)(A + BW)^{-1}]\}^{-1}.$$

Из обеих фигурных скобок вынесем вправо сомножитель $(A + BW)^{-1}$. Используя тот факт, что обратная матрица к произведению равна произведению обратных матриц в обратном порядке, получаем

$$\begin{aligned} & [M(A + BW) + N(C + DW)](A + BW)^{-1}(A + BW)[K(A + BW) + L(C + DW)]^{-1} \\ & = [(MA + NC) + (MB + ND)W][(KA + LC) + (KB + LD)W]^{-1}. \end{aligned}$$

Мы показали, что композиция преобразований Мебиуса приводит к преобразованию Мебиуса, отвечающему произведению матриц. Следовательно, группа преобразований Мебиуса определяет представление группы невырожденных матриц $\text{GL}(n)$, как группы преобразований многообразия Грассмана. При умножении матрицы на скаляр преобразование Мебиуса не меняется. Поэтому можно считать, что на $\text{Gr}(k, n)$ действует специальная линейная группа $\text{SL}(n)$ (группа матриц с определителем единица).

С этим действием естественно связано инвариантное описание многообразий Грассмана. Рассмотрим группу Ли ортогональных преобразований $\mathbb{O}(n)$ пространства \mathbb{R}^n , которая как подгруппа $\text{GL}(n)$ также действует на $\text{Gr}(k, n)$ и притом транзитивно, поскольку любую пару ортогональных плоскостей можно перевести в любую другую ортогональным преобразованием. Рассмотрим стационарную подгруппу плоскости W , т.е. подгруппу $\mathbb{O}(n)$, оставляющую эту плоскость неподвижной. Такое преобразование автоматически оставляет неподвижной и ортогональное дополнение W^\perp к плоскости W , поэтому в базисе, состоящем из векторов, принадлежащих W и W^\perp , стационарная подгруппа представляется блочно диагональными матрицами

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in (\mathbb{O}(k) \times \mathbb{O}(n - k)).$$

Поэтому грассманово многообразие естественно рассматривать как (для определенности левое) фактор пространство

$$\text{Gr}(k, n) = \mathbb{O}(n)/(\mathbb{O}(k) \times \mathbb{O}(n - k)). \quad (50)$$

Ортогональная группа (а не группа $\text{SL}(n)$) выбрана потому, что она является полупростой компактной группой Ли, и поэтому на ней определена структура евклидова пространства, задаваемая формой Киллинга на алгебре Ли, (т.е. на касательной плоскости в единице)

$(g_1, g_2) = -\text{tr}(\text{ad } g_1 \circ \text{ad } g_2)$. Действие группы Ли на фактор пространстве переносит эту метрику с касательной плоскости в единице на остальные касательные плоскости, и, тем самым, превращает $\text{Gr}(k, n)$ в риманово многообразие, на котором $\mathbb{O}(n)$ действует как группа движений. Из формулы (74) следует, что $\text{Gr}(k, n)$ изоморфна $\text{Gr}(n - k, n)$. Изоморфизм осуществляется взятием ортогонального дополнения.

Многообразия $\text{Gr}(k, n)$ не только риманово, но и симметрическое. Пусть $o \in \text{Gr}(k, n)$ отвечает единице группы $\mathbb{O}(n)$; X плоскость в \mathbb{R}^n , отвечающая этой точке и $Y = X^\perp$. Рассмотрим преобразование \mathbb{R}^n , задаваемое матрицей

$$\sigma = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}.$$

Это отображение тождественно на X , а матрица

$$V = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

переходит в матрицу

$$\sigma V \sigma^{-1} = \begin{pmatrix} A & -B \\ -C & D \end{pmatrix}.$$

Неподвижными остаются блочно диагональные матрицы и только они, т.е. элементы стационарной подгруппы. Получаем отображение $\text{Gr}(k, n) \rightarrow \text{Gr}(k, n)$ с единственной неподвижной точкой, задающее геодезическую симметрию. Множество таких отображений порождает группу изометрий и превращает $\text{Gr}(k, n)$ в симметрическое пространство.

Альтернативный способ задания координат на многообразии $\text{Gr}(k, n)$ это плюккеровы координаты — координаты поливекторов, отвечающих подпространствам W . Пусть $\{c_1, c_2 \dots c_k\}$ базис подпространства W . Простым поливектором $c_1 \wedge c_2 \wedge \dots \wedge c_k$ называется кососимметрический тензор, полученный альтернированием тензорного произведения $c_1 \otimes c_2 \otimes \dots \otimes c_k$. Операция альтернирования состоит в следующем: для каждой компоненты тензорного произведения, отвечающей мультииндексу $K = \{i_1, i_2 \dots i_k\}$, производится суммирование по всем перестановкам индексов, входящим в мультииндекс K , взятым со знаком плюс или минус в зависимости от четности перестановки. Такое суммирование дает детерминант, и координатами подпространства W является набор из всех детерминантов

$$C^{i_1, i_2 \dots i_k} = \begin{vmatrix} c_1^{i_1} & \dots & c_1^{i_k} \\ \vdots & \ddots & \vdots \\ c_k^{i_1} & \dots & c_k^{i_k} \end{vmatrix},$$

Это означает, что надо рассмотреть все миноры k -го порядка $(k \times n)$ -матрицы, составленной из векторов базиса подпространстве W . В результате получается набор из $\binom{n}{k}$ величин $C^{i_1, i_2 \dots i_k}$. Посмотрим, что происходит с этим набором при замене базиса $\{c_1, c_2 \dots c_k\}$ подпространства W на другой базис $\{b_1, b_2 \dots b_k\}$. Имеет место очевидная

Лемма 1. При замене базиса все величины $C^{i_1, i_2 \dots i_k}$ умножаются на один и тот же множитель — на якобиан перехода от базиса $\{c_1, c_2 \dots c_k\}$ к базису $\{b_1, b_2 \dots b_k\}$.

Тем самым плюккеровы координаты $C^{i_1, i_2 \dots i_k}$ плоскости W определяют точку $\left(\binom{n}{k} - 1\right)$ -мерного проективного пространства $\text{PR}^{\binom{n}{k}-1}$. Линейное пространство Λ^k , натянутое на простые поливекторы размерности k , называется k -ой внешней степенью исходного пространства \mathbb{R}^n .

Плюккеровы координаты k -мерных подпространств определяют отображение \mathcal{P} из множества этих подпространств в Λ^k . Из сравнения размерностей ясно, что отображение \mathcal{P} не может быть сюръективным. Миноры C^{i_1, i_2, \dots, i_k} зависимы. В силу того, что $\text{Gr}(k, n) \simeq \text{Gr}(n - k, n)$ можно без ограничения общности считать, что $k < n/2$ (за координаты плоскости можно принять координаты ее ортогонального дополнения). Для описания соотношений между минорами составим удвоенную $(2k \times n)$ -матрицу, полученную путем приписывания к k строкам базиса еще k таких же строк. Все $2k \times 2k$ -миноры полученной матрицы нулевые. Выберем любой $2k \times 2k$ -минор и разложим его по формуле Лапласа по первым k строкам. Получим квадратичное соотношение на плюккеровы координаты. Таких соотношений будет столько, сколько можно составить различных $2k \times 2k$ -миноров, т.е. $\binom{n}{2k}$. Эти квадратичные соотношения определяют в пространстве Λ^k конус \mathcal{K} , называемый плюккеровым конусом. В разделах о минимизации кратных интегралов нам понадобится следующее

Утверждение 1. *Рассмотрим автоморфизм $V : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Индуцированный этим отображением автоморфизм $\Lambda^k \rightarrow \Lambda^k$ задается матрицей $\Lambda^k V$, составленный из миноров порядка k матрицы V .*

Пример 3. *Рассмотрим $\text{Gr}(2, 4)$ — многообразие двумерных подпространств в \mathbb{R}^4 . Базис двумерной плоскости задается парой векторов $\{a_1, a_2, a_3, a_4\}, \{b_1, b_2, b_3, b_4\}$. Миноры второго порядка матрицы, составленной из этих векторов, обозначим через $\{A_{12}, A_{13}, A_{14}, A_{23}, A_{24}, A_{34}\}$. Разлагая удвоенную матрицу из четырех строк по первым двум строкам, получим соотношение $2(A_{12}A_{34} - A_{13}A_{24} + A_{14}A_{23}) = 0$. Поворот на угол $\pi/4$ в каждой из трех двумерных плоскостей пространства с координатами $\{A_{ij}\}$ превращает произведения в разность квадратов. В результате получаем уравнение*

$$x_1^2 + x_3^2 + x_5^2 = x_2^2 + x_4^2 + x_6^2$$

В силу однородности плюккеровых координат имеем уравнение прямого произведения двух двумерных сфер

$$\text{Gr}(2, 4) \simeq S^2 \times S^2.$$

Уравнение Эйлера для второй вариации (90) имеет вид

$$-\frac{d}{dt}(A\dot{h} + C^*h) + (C\dot{h} + Bh) = 0.$$

Запишем его в канонической форме. $p = A\dot{h} + C^*h$.

$$\begin{cases} \dot{h} = A^{-1}p - A^{-1}C^*h \\ \dot{p} = C(A^{-1}p - A^{-1}C^*h) + Bh = CA^{-1}p + (B - CA^{-1}C^*)h. \end{cases} \quad (51)$$

Это уравнение в вариациях для уравнения Эйлера (89), поэтому оно определяет перенос касательных векторов по решениям уравнения Эйлера. Лагранжевы многообразия это n -мерные поверхности, поэтому нас будет интересовать перенос n -мерных плоскостей вдоль линейной системы (75). Примем за базисные плоскости пространство векторов $\{h\}$ (горизонтальная плоскость) и пространство векторов $\{p\}$ (вертикальная плоскость). В соответствии с формулой (92) координатами n -мерной плоскости W служит матрица $W = ph^{-1}$ (в уравнении (75) p и h можно рассматривать как матрицы). Дифференцируя W в силу системы (75), получаем

$$\dot{W} = [CA^{-1}p + (B - CA^{-1}C^*)h]h^{-1} - ph^{-1}(A^{-1}p - A^{-1}C^*h)h^{-1}$$

или

$$\dot{W} + (C - W)A^{-1}(C^* - W) - B = 0,$$

что совпадает с уравнением (91). В рассматриваемом случае уравнение Риккати имеет еще одну дополнительную симметрию — симплектическую структуру.

14 ТЕОРЕМА ЯКОБИ ОБ ОГИБАЮЩЕЙ

Типичная ситуация возникновения сопряженной точки на экстремали — это ее касание с огибающей однопараметрического семейства экстремалей, исходящих из одной точки. Такие огибающие были найдены Якоби при исследовании геодезического потока на трехосном эллипсоиде.

Приведем здесь теорему, характеризующую замечательное геометрическое свойство огибающих однопараметрического семейства экстремалей. При этом вместо задачи с фиксированными концами, когда все экстремали выходят из одной точки, будет рассмотрена более общая задача с подвижным левым концом, что, практически, не усложнит выкладок.

Рассмотрим функционал

$$\mathcal{J} = \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t)) dt \quad (52)$$

с граничным условием $\Phi(t_0, x(t_0)) = 0$. Здесь $x \in \mathbb{R}$, $f : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$; $f \in C^1(D)$; $\Phi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $\Phi \in C^1(D)$, $\Phi' \neq 0$; D — окрестность графика кривой $x = \hat{x}(t)$, $\dot{x} = \hat{\dot{x}}(t)$, $t_0 \in [\hat{t}_0, \hat{t}_1]$, где $\hat{x}(\cdot)$ — экстремаль функционала \mathcal{J} , удовлетворяющая условию трансверсальности (26) в точке $(\hat{t}_0, \hat{x}(\hat{t}_0))$. Рассмотрим однопараметрическое семейство экстремалей $x(t, \alpha)$, удовлетворяющих условию трансверсальности (26), и такое, что $x(t, 0) = \hat{x}(t)$. Предположим, что это семейство имеет огибающую (дуга BD на рис. 5); AC на этом рисунке — это кривая $\Phi(t, x) = 0$; AB и CD — экстремали, соответствующие значениям α_1 и α_2 . Момент касания экстремали $x(t, \alpha)$ с огибающей обозначим через $t_1(\alpha)$.

Теорема 4 (Теорема Якоби об огибающей). .

Если параметризация огибающей BD выбрана так, что скорость движения по BD в каждой точке равна скорости движения по экстремали, касающейся BD в этой точке, то разность значений функционала \mathcal{J} на экстремальных AB и CD равна значению \mathcal{J} на отрезке огибающей BD .

Доказательство.

Подставив $x(t, \alpha)$ под знак функционала \mathcal{J} , получим скалярную функцию $\mathcal{J}(\alpha)$. В силу формулы (23)

$$\frac{d\mathcal{J}}{d\alpha} = \int_{t_0(\alpha)}^{t_1(\alpha)} [f_x - \frac{d}{dt} f_{\dot{x}}] \frac{\partial x}{\partial \alpha} dt + [p\delta x - \mathbb{H}\delta t]_{t_0(\alpha)}^{t_1(\alpha)}.$$

Интегральный член в этой формуле равен нулю, поскольку $x(t, \alpha)$ — экстремаль, $[p\delta x - \mathbb{H}\delta t]_{t_0(\alpha)} = 0$ в силу выполнения условия трансверсальности (25) в точках $(t_0(\alpha), x(t_0(\alpha), \alpha))$. Поэтому

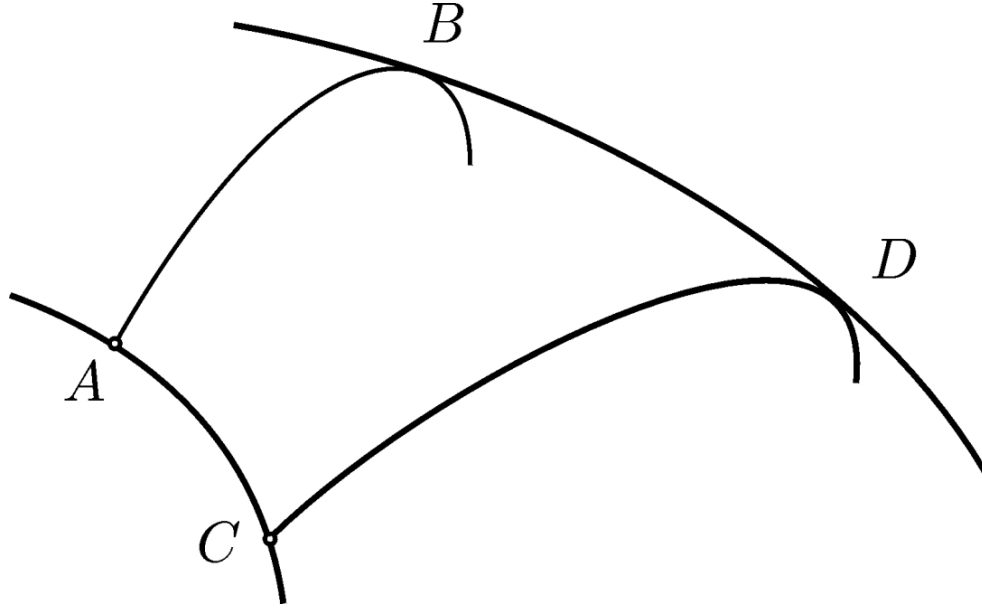


Рис. 5:

$$\frac{d\mathcal{J}}{d\alpha} = f_x \frac{\partial x}{\partial \alpha}(t_1(\alpha), \alpha) - [-f + f_x \dot{x}] \frac{dt_1}{d\alpha}. \quad (53)$$

Поскольку BD – огибающая, касательная к ней, задаваемая вектором $(\frac{dt_1}{d\alpha}, \frac{dx}{d\alpha}(t_1(\alpha), \alpha))$, совпадает с касательной к экстремали, задаваемой вектором $(1, \dot{x}(t_1(\alpha), \alpha))$, т. е. $\dot{x}(t_1(\alpha), \alpha) \frac{dt_1}{d\alpha} = \frac{\partial x}{\partial \alpha}(t_1(\alpha), \alpha)$. Подставив это соотношение в (53), получим

$$\frac{d\mathcal{J}}{d\alpha} = f(t_1(\alpha), x(t_1(\alpha), \alpha), \dot{x}(t_1(\alpha), \alpha)). \quad (54)$$

Интегрируя (54) от α_1 до α_2 имеем

$$\mathcal{J}(\alpha_2) - \mathcal{J}(\alpha_1) = \int_{\alpha_1}^{\alpha_2} f(t_1(\alpha), x(t_1(\alpha), \alpha), \dot{x}(t_1(\alpha), \alpha)) d\alpha. \quad (55)$$

Правая часть формулы (55) есть значение функционала \mathcal{J} на BD , о котором говорится в условиях теоремы.

□

Доказанная теорема обобщает известное свойство эволюты плоской кривой (см. [?] с.129). Напомним, что эволютой плоской кривой $y = \varphi(x)$ называется геометрическое место ее центров кривизны, и что эволюта является огибающей семейства нормалей к исходной кривой $y = \varphi(x)$. На множестве плоских кривых $x = x(t)$, $y = y(t)$, $t_0 \leq t \leq t_1$, удовлетворяющих граничным условиям $y(t_0) = \varphi(x(t_0))$ рассмотрим функционал $\int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2} dt$, определяющий длину кривой. Экстремалиями для него служат прямые; условию трансверсальности удовлетворяют нормали. Огибающей семейства нормалей будет эволюта. Применим теорему Якоби.

Если в двух точках A и C кривой $y = \varphi(x)$ (см. рис. 6) провести нормали (AB и CD соответственно) и рассмотреть расстояния по ним до точек B и D касания с эволютой, то величина $(|AB| - |CD|)$ будет равна длине дуги эволюты BD .

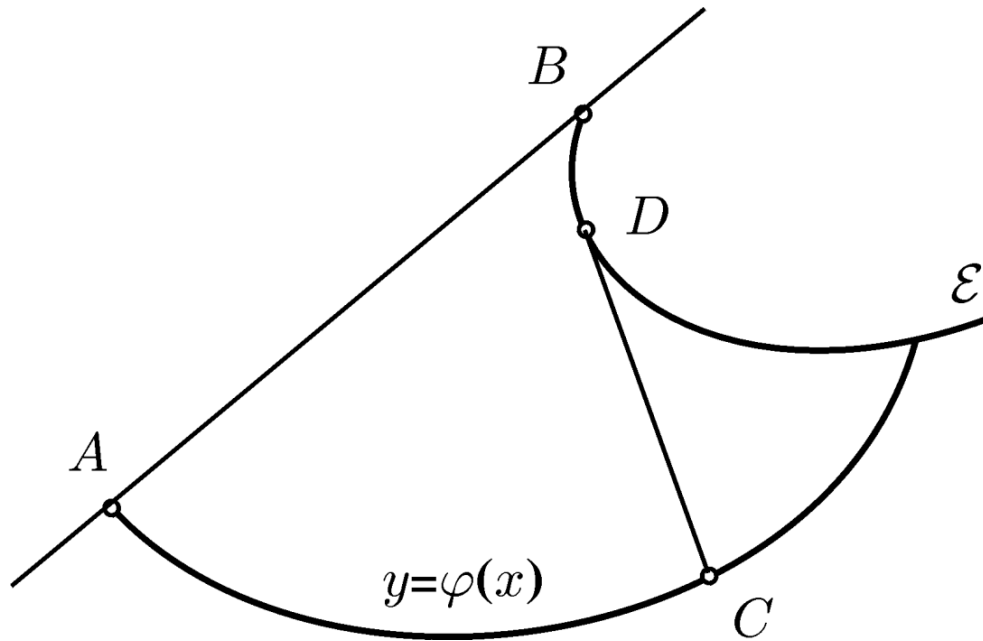


Рис. 6:

Доказанный факт имеет изящное механическое истолкование. Представим себе гибкую нерастяжимую нить закрепленную в точке B и накрученную на кривую BDE . Если эту нить начать разворачивать, сматывая с эволюты, но сохраняя ее в натянутом состоянии, то конец E опишет кривую $y = \varphi(x)$.

Пример 4. .

На множестве плоских кривых $x = x(t)$, $y = y(t)$, $0 \leq t \leq 1$, удовлетворяющих граничным условиям $x(0) = \xi$, $y(0) = \eta$, $y(1) = x^2(1)$, рассмотрим функционал $\int_0^1 \sqrt{\dot{x}^2 + \dot{y}^2} dt$, определяющий длину кривой. Задача минимизации этого функционала соответствует нахождению экстремального расстояния от точек плоскости (ξ, η) до параболы $y = x^2$.

Эта задача является одним из вариантов задачи Аполония, который, изучая расстояние до эллипса, фактически нашел его эволюту. Обобщением этой задачи является задача нахождения на римановом многообразии экстремалей расстояния до данного подмногообразия, которая послужила основой для создания теории Морса.

Как и в вышеприведенном случае, экстремалами, удовлетворяющими условию трансверсальности, служат нормали. Уравнение нормали в точке (a, a^2) имеет вид $y - a^2 = -\frac{1}{2a}(x - a)$. Огибающая этого семейства удовлетворяет системе $2ay - 2a^3 = a - x$, $2y - 6a^2 = 1$. Исключая параметр a , получаем уравнение полукубической параболы (см. рис. 7). В некоторой фиксированной точке A параболы $y = x^2$ рассмотрим нормаль, касающуюся эволюты в точке B . Тогда $AB = R$ есть радиус кривизны параболы в точке A . Если точка (ξ, η) лежит на этой нормали между точками A и B (точка B_1 на рис. 7), то окружность с центром в точке B_1 радиуса $|AB_1|$ имеет кривизну большую, чем кривизна параболы в точке A . Следовательно эта окружность касается параболы "изнутри т. е. в окрестности точки A она лежит в области и $y \geq x^2$. Тем самым, на AB_1 достигается локальный минимум.

Если $|AB_2| > R$, то окружность с центром в точке B_2 радиуса $|AB_2|$ имеет кривизну, меньшую чем кривизна параболы в точке A и касается параболы "снаружи т. е. в окрестности

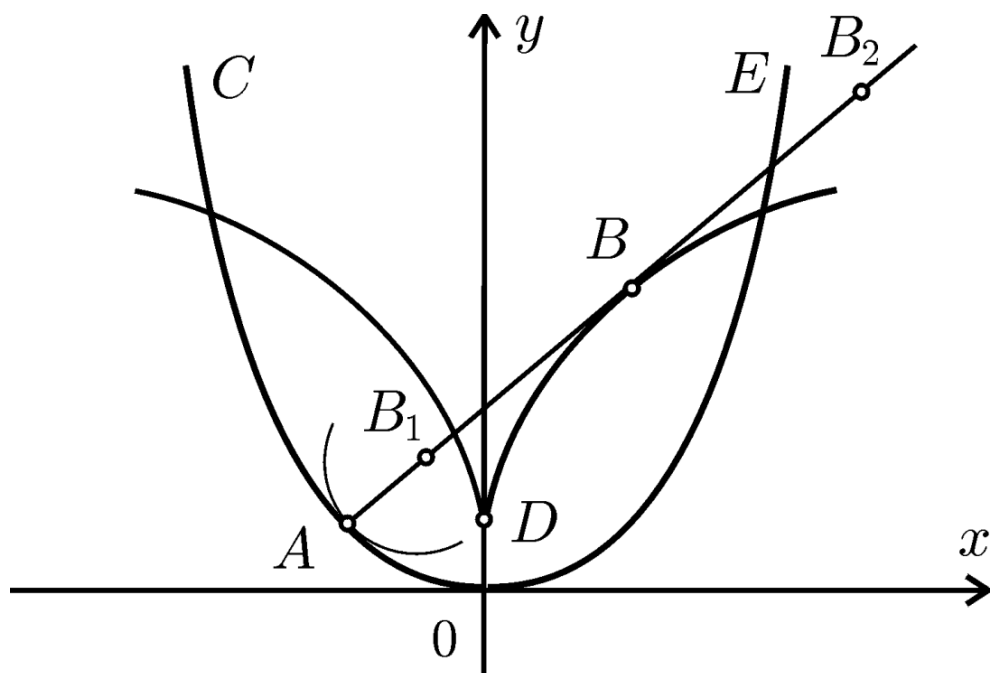


Рис. 7:

точки A лежит в области $y \leq x^2$. На AB_2 достигается локальный максимум. Для точки B соприкасающаяся окружность переходит с одной стороны параболы на другую – экстремума нет.

Таким образом, огибающая однопараметрического семейства нормалей (дуга CDE на рис. 7) играет в примере 19 ту же роль, что и сопряженные точки в задачах с фиксированными концами: до точки касания с огибающей экстремаль локально минимальна, после касания она перестает быть локально минимальной. Точки кривой CDE называются фокальными точками параболы.

Для того, чтобы пояснить смысл этого понятия, напомним наглядное определение огибающей однопараметрического семейства прямых на плоскости: "огибающая есть геометрическое место точек пересечения бесконечно близких прямых". Иными словами, если одна из прямых "стремится" к другой, то точка их пересечения стремится к точке, лежащей на огибающей.

Если представить себе лучи света, испускаемые раскаленной нитью, имеющей форму параболы (речь идет о плоской задаче распространения света в однородной изотропной среде), то точки кривой CDE окажутся наиболее ярко освещенными. Если среда неоднородна и неизотропна, то роль прямых будут играть экстремали соответствующего функционала. Если размерность пространства больше двух, то вместо пересечения надо рассматривать сближение. Действительно, касательные к пространственной кривой (которая является для них огибающей) не пересекаются между собой, но при этом точка касания с огибающей – это предельное положение точек максимального сближения бесконечно близких прямых.

С этой точки зрения между понятиями "сопряженная точка" и "фокальная точка" принципиальной разницы нет. В первом случае семейство экстремалей исходит из одной общей точки по всем направлениям (его особенности – это сопряженные точки); во втором – из точек многообразия по направлениям, удовлетворяющим условию трансверсальности (его особенности – это фокальные точки).

Понятие "фокальная точка" по смыслу близко понятию "каустика используемому в теории дифференциальных уравнений в частных производных. Кстати сказать, оба названия имеют общую этимологию: focus (фокус) – по латыни – очаг, жар; *καυστικός* (каустикос) – по гречески – воспламеняющий, обжигающий (ср. русское "каустическая сода"). Эти названия связаны с оптикой и характеризуют точки пространства, в которых происходит увеличение плотности энергии светового потока, вызванное сближением лучей света.

15 НЕОБХОДИМОЕ УСЛОВИЕ ВЕЙЕРШТРАССА

Определение 1. Будем говорить, что кривая $\hat{x}(\cdot)$ реализует сильный минимум в задаче 5 (гл. V § 17), если существует такая окрестность V точки $\hat{x}(\cdot)$ в пространстве $\mathcal{K}C^1([t_0, t_1])$, что для любой кривой $x(\cdot) \in V$ $\mathcal{J}(x(\cdot)) \geq \mathcal{J}(\hat{x}(\cdot))$.

Очевидно, что всякий сильный минимум является слабым, и поэтому достаточные условия сильного минимума являются одновременно достаточными условиями слабого, а необходимые условия слабого минимума – необходимыми условиями сильного. Однако не всякий слабый минимум является сильным (см. дальше пример 20). Поэтому естественно поставить вопрос о необходимых условиях сильного минимума, не являющихся необходимыми для слабого. Именно таким является условие Вейерштрасса. Выведем это условие.

Фиксируем $\tau \in [t_0, t_1]$, вектор $\xi \in \mathbb{R}^n$ и число $\sigma > 0$. Рассмотрим семейство кривых (см. рис. 8).

$$x(t, \alpha) = \begin{cases} \hat{x}(t), & t \notin [\tau, \tau + \sigma] \\ \hat{x}(\tau) + \xi(t - \tau), & t \in [\tau, \tau + \alpha] \\ \hat{x}(t) + \frac{\hat{x}(\tau) + \xi\alpha - \hat{x}(\tau + \alpha)}{\sigma - \alpha}(\tau + \sigma - t), & t \in [\tau + \alpha, \tau + \sigma] \end{cases}$$

При достаточно малом α кривые $x(t, \alpha)$ попадают в сколь угодно малую C -окрестность (но не C^1 -окрестность!) кривой $\hat{x}(\cdot)$; $x(t, 0) = \hat{x}(t)$.

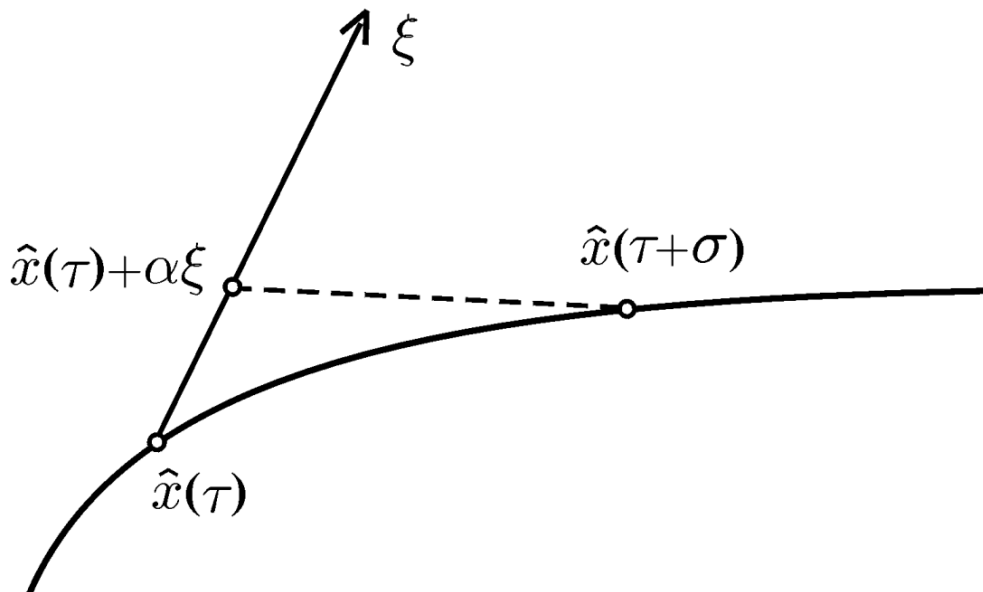


Рис. 8:

Подставив $x(t, \alpha)$ под знак функционала \mathcal{J} , получим функцию $\mathcal{J}(\alpha)$, которая определена при $\alpha \geq 0$ и достигает минимума при $\alpha = 0$. Поэтому $\frac{d\mathcal{J}}{d\alpha} \geq 0$. Имеем $\mathcal{J}(\alpha) - \mathcal{J}(0) = \int_{\tau}^{\tau+\alpha} f(t, \hat{x}(t) + \xi(t - \tau), \xi) dt - \int_{\tau}^{\tau+\sigma} f(t, \hat{x}, \dot{\hat{x}}) dt + \int_{\tau+\alpha}^{\tau+\sigma} f(t, x(t, \alpha), \dot{x}(t, \alpha)) dt$. Применив формулу (??), получим

$$\frac{d\mathcal{J}}{d\alpha} = t(\tau + \alpha, \hat{x}(\tau) + \xi\alpha, \xi) - [p\delta x - \mathbb{H}\delta t]_{\tau+\alpha} + \int_{\tau-\alpha}^{\tau+\alpha} [f_x - \frac{d}{dt}f_{\dot{x}}] \frac{\partial x}{\partial \alpha} dt. \quad (56)$$

Поскольку на $\hat{x}(\cdot)$ выполнено уравнение Эйлера, интегральный член при $\alpha = 0$ дает ноль. Так как левый конец семейства $x(t, \alpha)$ скользит по прямой $t = \tau + \alpha$, $x = \hat{x}(\tau) + \alpha\xi$ с направляющим вектором $(1, \xi)$, вместо вектора $(\delta t, \delta x)$ надо подставить вектор $(1, \xi)$. В результате получим

$$f(\tau, \hat{x}(\tau), \xi) - f(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau)) - f_{\dot{x}}(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau))(\xi - \dot{\hat{x}}(\tau)) \geq 0. \quad (57)$$

Функция $\mathcal{E}(\tau, x, \dot{x}, \xi) = f(\tau, x, \xi) - f(\tau, x, \dot{x}) - f_{\dot{x}}(\tau, x, \dot{x})(\xi - \dot{x})$ называется функцией Вейерштрасса. Нами доказана следующая

Теорема 5. Пусть экстремаль $\hat{x}(\cdot)$ дает сильный минимум функционалу \mathcal{J} . Тогда $\mathcal{E}(\tau, \hat{x}(\tau), \dot{\hat{x}}(\tau), \xi) \geq 0$ при всех $\xi \in \mathbb{R}^n$, $\tau \in [t_0, t_1]$.

Обсудим смысл этой теоремы. Зафиксируем значения τ_0, x_0, \dot{x}_0 и рассмотрим $y = f(\tau_0, x_0, \xi)$ как функцию от переменного ξ . Тогда функцию Вейерштрасса можно интерпретировать как уклонение поверхности $y = f(\tau_0, x_0, \xi)$ от касательной плоскости, проведенной в точке \dot{x}_0 (см. рис. 9). Тем самым, условие Вейерштрасса означает, что при всех $\tau \in [t_0, t_1]$ поверхность $y = f(\tau, \hat{x}(\tau), \xi)$ целиком лежит над касательной плоскостью, проведенной в точке $\dot{\hat{x}}(\tau)$.

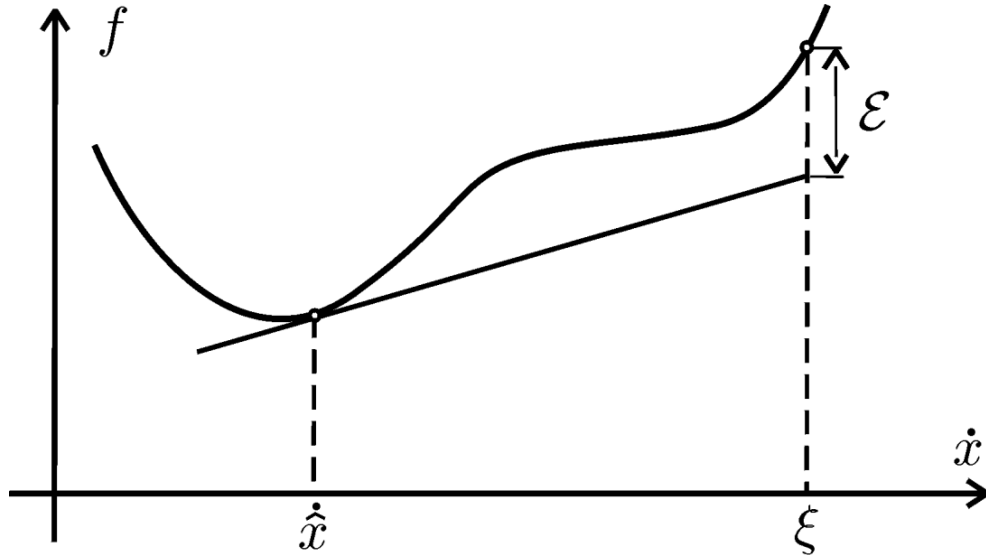


Рис. 9:

Замечание 2.

Применив формулу Тейлора

$$f(t, \hat{x}(t), \xi) = f(t, \hat{x}(t), \dot{\hat{x}}(t)) + f_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))(\xi - \dot{\hat{x}}(t)) + \frac{1}{2} \langle f_{\dot{x}\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t) + \theta(\xi - \dot{\hat{x}}(t))) (\xi - \dot{\hat{x}}(t)), (\xi - \dot{\hat{x}}(t)) \rangle,$$

где $0 \leq \theta \leq 1$, легко показать, что из условия Вейерштрасса следует условие Лежандра $\hat{f}_{\dot{x}\dot{x}} \geq 0$ (но не следует усиленное условие Лежандра).

16 ИНТЕГРАЛЬНЫЙ ИНВАРИАНТ ПУАНКАРЕ-КАРТАНА

Идея интегрального инварианта возникла в работах Стокса при изучении стационарного вихревого течения несжимаемой жидкости. Приведем рассуждения Стокса, следуя изложению в ([?] с.200).

Пусть v – векторное поле скоростей течения жидкости; $\text{rot } v$ – поле ротора. Интегральные кривые векторного поля ротора называются вихревыми линиями. Рассмотрим замкнутый одномерный контур γ_1 и через каждую его точку проведем вихревую линию. Получим вихревую трубку. Рассмотрим контур γ_2 , охватывающий ту же вихревую трубку (см. рис. 10). Боковую поверхность этой трубки с границей $\gamma_1 - \gamma_2$ обозначим через Γ .

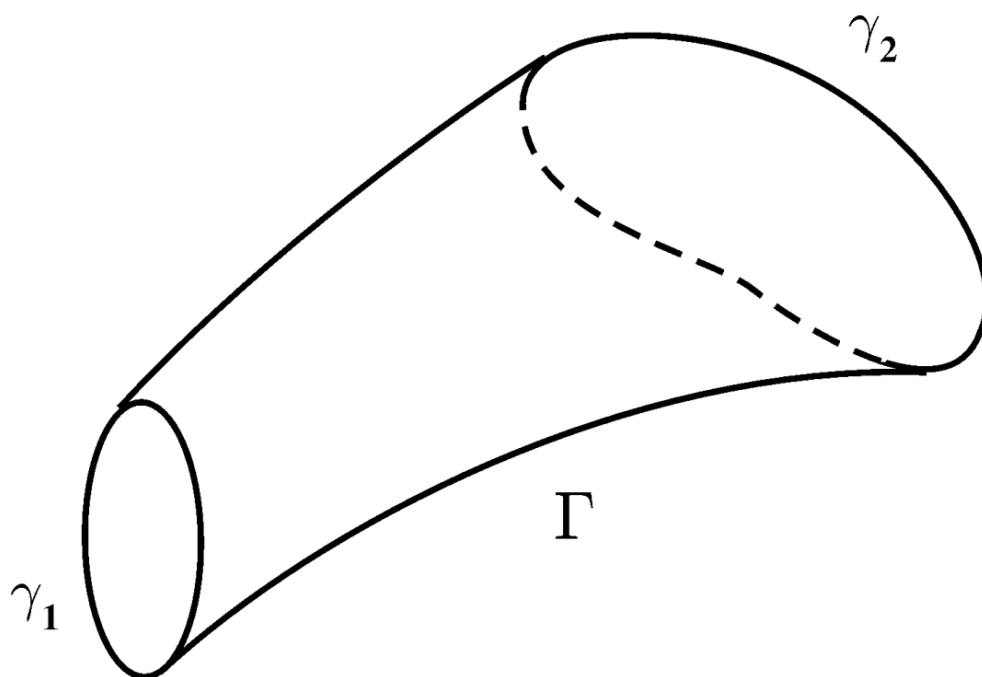


Рис. 10:

Циркуляцией поля v по контуру γ называется $\int_{\gamma} v dl$, где форма $v dl$ имеет вид $v dl = v_1 dx_1 + v_2 dx_2 + v_3 dx_3$.

Лемма 2. (Стокс).

Циркуляции поля v по γ_1 и γ_2 одинаковы, т. е. $\int_{\gamma_1} v dl = \int_{\gamma_2} v dl$.

Доказательство.

Применяя к Γ формулу Стокса, имеем $\int_{\gamma_1} v dl - \int_{\gamma_2} v dl = \int_{\Gamma} (\text{rot } v)_n dS = 0$. Последнее равенство есть следствие того, что $\text{rot } v$ всюду касается Γ и его поток через Γ равен нулю.

□

Утверждение леммы Стокса можно выразить следующей фразой: форма $v dl$ есть интегральный инвариант для поля ротора.

Впоследствии А. Пуанкаре показал, что если рассмотреть отображение $\mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$, $(x_0, p_0) \mapsto x(t; x_0, p_0), p(t; x_0, p_0)$ по траекториям канонической системы $\dot{x} = Hp$, $\dot{p} = -Hx$, то дифференциальная форма $p dx$ есть интегральный инвариант этого отображения. Дальнейшее обобщение этой идеи принадлежит Э. Картану. В последующем изложении роль циркуляции играет интеграл по замкнутому контуру от произвольной 1-формы $\omega^{(1)}$; роль ротора – билинейная форма $\omega^{(2)} = d\omega^{(1)}$. Ранг билинейной кососимметрической формы – четное число. Дифференциальная 2-форма $\omega^{(2)}$ называется невырожденной, если ее ранг в каждой точке максимален. В нечетномерном пространстве \mathbb{R}^{2k+1} невырожденные 2-формы имеют ранг $2k$. В этом случае матрица отображения A_x , определяемого формой $\omega^{(2)}$, имеет одномерное ядро. Обозначим через $\zeta(x)$ векторное поле, направленное по этому ядру.

Если двумерная плоскость σ , натянутая на вектора ξ, η , содержит $\zeta(x)$, то $\omega_x^{(2)}\{\xi, \eta\} = 0$. Действительно, один из векторов, скажем, ξ , есть линейная комбинация ζ и η . Поэтому $\omega_x^{(2)}\{\xi, \eta\} = \omega_x^{(2)}\{\alpha\zeta + \beta\eta, \eta\} = \alpha\langle A_x\zeta, \eta \rangle + \beta\omega_x^{(2)}\{\eta, \eta\}$. Первый член в правой части равен нулю, поскольку ζ принадлежит ядру A , а второй – в силу кососимметричности $\omega^{(2)}$.

Пусть теперь $\omega^{(1)}$ такая 1-форма в \mathbb{R}^{2k+1} , что ее дифференциал $\omega^{(2)} = d\omega^{(1)}$ невырожден. Интегральные линии векторного поля $\zeta(x)$ называются характеристиками формы $\omega^{(1)}$. Для одномерного контура γ_1 рассмотрим трубку из характеристик, проходящих через γ_1 , и контур γ_2 , охватывающий эту трубку (рис. 10). Применяя формулу Стокса к двумерному многообразию с границей $\gamma_1 - \gamma_2$, получим

$$\int_{\gamma_1} \omega^{(1)} - \int_{\gamma_2} \omega^{(1)} = \int_{\Gamma} d\omega^{(1)} = 0.$$

Действительно, форма $d\omega_x^{(1)}$ на 2-мерной площадке σ_x , касательной к Γ в точке x , обращается в 0, т. к. σ_x содержит вектор ζ_x . Таким образом, можно сказать, что форма $\omega^{(1)}$ есть интегральный инвариант для поля своих характеристических направлений.

Теперь мы можем перейти к каноническим системам. Рассмотрим гамильтониан $H(t, x, p)$, определенный на пространстве переменных (t, x, p) , которое является нечетномерным ($t \in \mathbb{R}$, $(x, p) \in \mathbb{R}^{2n}$) и дифференциальную форму Пуанкаре-Картана $\omega^{(1)} = p dx - H dt$. Применим к этой форме вышеописанную конструкцию. Матрица A формы $\omega^{(2)} = d\omega^{(1)}$ имеет вид:

$$\begin{pmatrix} & & & & & & -\partial H/\partial x_1 & \\ & & & & & & \vdots & \\ & & & & & & -\partial H/\partial x_n & \\ & & & & & & -\partial H/\partial p_1 & \\ & & & & & & \vdots & \\ & & & & & & -\partial H/\partial p_n & \\ \partial H/\partial x_1 & \cdots & \partial H/\partial x_n & \partial H/\partial p_1 & \cdots & \partial H/\partial p_n & 0 & \end{pmatrix}$$

Эта матрица при всех (t, x, p) имеет максимальный ранг $2n$, т. к. ее левый верхний угол – невырожденная $(2n \times 2n)$ -матрица. Определенный выше вектор $\zeta(x)$ для матрицы A имеет отличную от нуля последнюю компоненту и может быть нормирован условием $\zeta_{2n+1} = 1$. Тогда

получаем $-\zeta_{n+1} - \partial H/\partial x_1 = 0, \dots - \zeta_{2n} - \partial H/\partial x_n = 0, \zeta_1 - \partial H/\partial p_1 = 0, \dots \zeta_n - \partial H/\partial p_n = 0, \zeta_{2n+1} = 1$, т. е. уравнения характеристик формы Пуанкаре-Картана

$$\frac{dp}{ds} = -\frac{\partial H}{\partial x}, \quad \frac{dx}{ds} = \frac{\partial H}{\partial p}, \quad \frac{dt}{ds} = 1. \quad (58)$$

дают каноническую систему с гамильтонианом H . Тем самым, доказана следующая

Теорема 6. *(Об интегральном инварианте).*

Дифференциальная форма $pdx - Hdt$ является интегральным инвариантом канонической системы (58).

Это означает, что интеграл формы $pdx - Hdt$ по любому замкнутому контуру сохраняется вдоль трубки траекторий системы (58).

17 ОСНОВНАЯ КОНСТРУКЦИЯ

Пусть дана форма Пуанкаре-Картана $pdx - Hdt = \omega^{(1)}$, $H \in C^2(D)$, где $D \subset \mathbb{R}^{2n+1}$ – открытое множество.

Определение 4.

Подмногообразие $M \subset D$ называется интегральным многообразием формы ω , если форма ω обращается в ноль на любых наборах векторов касательных к M : $\omega|_M = 0$.

Дифференциальная форма $\omega^{(1)} = pdx - Hdt$ подсказывает нам, что переменная p сопряжена с переменной x , а переменная H сопряжена с переменной t . Используя форму $\omega^{(1)}$, можно ввести симплектическую структуру в $(2n + 2)$ -мерное пространство переменных (x, p, t, H) . Для задач оптимального управления H является однозначной функцией от переменных (x, p, t) , т.е. любое многообразие пространства (x, p, t) допускает однозначный подъем в пространство (x, p, t, H) . Поэтому мы сохраним термин лагранжево многообразие для $(n + 1)$ -мерных подпространств $(2n + 1)$ -мерного пространства.

Определение 5.

Многообразие $M \subset D$, $\dim M = n + 1$ называется лагранжевым относительно формы $\omega^{(1)}$, если M – интегральное многообразие для формы $d\omega^{(1)}$.

Пример 5.

В каждой точке x многообразия $N \subset \mathbb{R}^n$, $\dim N = k$, рассмотрим множество векторов $p \in \mathbb{R}^n$, таких, что

$$\langle p, \xi \rangle = 0 \quad \text{для любых } \xi \in T_x N. \quad (59)$$

Множество элементов $\{x, p\}$, обладающих этим свойством, называется нормальным расслоением для многообразия N .

Нормальное расслоение для произвольного многообразия $N \subset \mathbb{R}^n$, очевидно, является интегральным многообразием формы $\omega^{(1)}$ при любом выборе функции H . Размерность нормального расслоения равна n , поскольку условие $\langle p, \xi \rangle = 0$ дает k независимых уравнений, если само N определяется системой из $(n - k)$ независимых уравнений.

Для многообразия $N \subset \mathbb{R}^{n+1}$, $\dim N = k$ обозначим через M множество элементов $\{t, x, p\}$, таких, что

$$\langle p, \xi \rangle - \mathbb{H}(t, x, p)\tau = 0 \quad \text{для любых } (t, x) \in N; (\tau, \xi) \in T_{(tx)}N. \quad (60)$$

Тогда M – интегральное многообразие формы $pdx - Hdt$.

Определение 6. .

Построенное таким образом многообразие M будем называть поднятием многообразия N в пространство t, x, p .

Рассмотрим задачу классического вариационного исчисления с подвижным левым и неподвижным правым концом:

Задача 2. .

$$\mathcal{J} = \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t))dt \rightarrow \inf \quad \text{при условиях } x(t_1) = x_1;$$

$$\Phi(t_0, x(t_0)) = 0, \quad \Phi : U \rightarrow \mathbb{R}^{n+1-k}, \quad \Phi \in C^1(U), \quad (61)$$

где $U \subset \mathbb{R}^1 \times \mathbb{R}^n$ – связная область, проекция которой на ось t содержит отрезок $[t_0, t_1]$.

Если отображение $\Phi' : U \rightarrow \mathbb{R}^{n+1-k}$ сюръективно, то условие (61) определяет в U k -мерное гладкое многообразие N . Рассмотрим многообразие M , построенное в примере .. по многообразию N и функции $H(t, x, p)$ (для задачи 1). В условии (60) нетрудно узнать условие трансверсальности для задачи 1.

Определение 7. .

Будем говорить, что многообразие M в точке (t, x, p) не содержит характеристических направлений, если вектор $\zeta = (1, H_p, -H_x)$ не касается многообразия M , т. е. $\zeta \notin TM$.

Предложение 2. .

Достаточным условием того, чтобы M в точке (t, x, p) не содержало характеристических направлений, является условие

$$\text{rk}(\xi_i - H_p(t, x, p)\tau_i) = k, \quad (62)$$

где $\{\tau_i, \xi_i\}$, $i = 1, \dots, k$ – базис $T_{tx}N$.

Доказательство. Достаточно показать, что вектор $(1, H_p)$ не касается многообразия N , т. е. не является линейной комбинацией векторов (τ_i, ξ_i) . Предположим противное: $1 = \sum_{i=1}^k \lambda_i \tau_i$, $H_p = \sum_{i=1}^k \lambda_i \xi_i$. Тогда $\sum_{i=1}^k \lambda_i \xi_i - H_p \sum_{i=1}^k \lambda_i \tau_i = H_p - H_p = 0$, что противоречит (62).
□

Следствие 1. .

Из вышеприведенного доказательства следует, что при выполнении условия (62) экстремали задачи 1, удовлетворяющие условию трансверсальности, подходят к многообразию N под ненулевым углом.

При выполнении условия (62) множество значений p , удовлетворяющих в точке (t, x) равенствам (??), образует гладкое многообразие. Обозначим это многообразие через $Q_{(tx)}$.

Основная конструкция.

Пусть M – поднятие многообразия N , определяемого формулой (59). Проведем через каждую точку (t_0, x_0, p_0) многообразия M экстремаль задачи 1, т. е. решение для $t \geq t_0$ гамильтоновой системы

$$\dot{x} = H_p(t, x, p), \quad \dot{p} = -H_x(t, x, p). \quad (63)$$

Если выполнено условие (62), то в результате получится $(n + 1)$ -мерное гладкое многообразие, которое мы обозначим через \mathcal{M} .

Предложение 3. . Если выполнено условие (62), то \mathcal{M} является лагранжевым многообразием относительно формы Пуанкаре-Картана.

Доказательство.

Рассмотрим произвольный гладкий замкнутый контур $\gamma \subset \mathcal{M}$ и проведем через него трубку траекторий Γ системы (63). Решения уравнений (63) гладко зависят от начальных данных и пересекают M под ненулевым углом. Поэтому $\gamma_1 = \Gamma \cap M$ – это гладкий контур. В силу теоремы об интегральном инварианте $\int_{\gamma} \omega^{(1)} = \int_{\gamma_1} \omega^{(1)}$. Но M – интегральное многообразие формы $\omega^{(1)}$. Поэтому

$$\int_{\gamma} \omega^{(1)} = 0 \quad \text{для любой замкнутой кривой } \gamma \subset \mathcal{M}. \quad (64)$$

Зафиксируем точку $(t_0, x_0, p_0) \in \mathcal{M}$ и рассмотрим $\int_{t_0 x_0 p_0}^{txp} \omega^{(1)}$ как функцию переменного верхнего предела $(t, x, p) \in \mathcal{M}$ (в силу формулы (64) интеграл не зависит от пути, если этот путь лежит в \mathcal{M}). Тогда дифференциал этой функции, определенной на \mathcal{M} , равен $\omega^{(1)}|_{\mathcal{M}}$. Тем самым, доказана точность формы $\omega^{(1)}|_{\mathcal{M}}$. Ее замкнутость есть следствие точности.

□

Следствие 2.

Предположим, что многообразие M связно. Рассмотрим две кривые l_1 и l_2 , соединяющие точку $(t, x, p) \in \mathcal{M}$ с многообразием M ; $l_1, l_2 \subset \mathcal{M}$. Тогда $\int_{l_1} \omega^{(1)} = \int_{l_2} \omega^{(1)}$.

Доказательство.

Соединив концы кривых l_1 и l_2 путем в M , получим замкнутый контур, интеграл по которому в силу формулы (64) равен нулю. Но M – интегральное многообразие формы $\omega^{(1)}$, поэтому интеграл по любой кривой, лежащей на M , равен нулю. Следовательно, $\int_{l_1} \omega^{(1)} = \int_{l_2} \omega^{(1)}$, как и утверждалось.

□

Всюду в дальнейшем под N, M, Q, \mathcal{M} будет пониматься результат применения основной конструкции к задаче 1.

18 ПОЛЕ ЭКСТРЕМАЛЕЙ

Рассмотрим в пространстве $\{t, x, p\}$ область Δ , содержащую наряду с любой точкой (t_1, x_1, p_1) все точки решения стандартной канонической системы с начальными значениями (t_1, x_1, p_1) , т. е. $(\tau, x(\tau; t_1, x_1, p_1), p(\tau; t_1, x_1, p_1))$, при $\tau < t_1$. Пусть $D = (\mathcal{M} \setminus M) \cap \Delta$.

Рассмотрим операцию проектирования $\pi : (t, x, p) \mapsto (t, x)$ из \mathbb{R}^{2n+1} в \mathbb{R}^{n+1} .

Определение 8.

Пусть проекция π , ограниченная на множество D является взаимно однозначным отображением. Тогда говорят, что в области $G = \pi D$ задано поле экстремалей \mathcal{P} с начальным многообразием N .

Если в G определено поле, то каждой точке $(t, x) \in G$ можно поставить в соответствие одну и только одну точку $p(t, x)$ так, что $(t, x, p(t, x)) \in D$. Иными словами, $p = p(t, x)$ – уравнение, определяющее многообразие \mathcal{M} , в области D .

В этой ситуации проекции экстремалей, лежащих в \mathcal{M} , на область G не пересекаются, поскольку экстремали не могут пересекаться в \mathcal{M} (по теореме единственности для канонической системы, а отображение $\pi|_{\mathcal{M}}$ взаимно однозначно. Поэтому в каждой точке $(t, x) \in G$ однозначно определена функция

$$g(t, x) = H_p(t, x, p(t, x)). \quad (65)$$

Она называется геодезическим наклоном поля \mathcal{P} . Уравнения $p = p(t, x)$ позволяют переписать все формулы, относящиеся к множеству D , в виде формул на G . Так, например, уравнения экстремалей в G имеют вид $\dot{x} = g(t, x)$; ограничение формы Пуанкаре-Картана на D можно записать в виде формы на G :

$$p(t, x)dx - H(t, x, p(t, x))dt. \quad (66)$$

Поскольку $\omega^{(1)}|_{\mathcal{M}}$ точна, форма (66) также является точной.

Замечание 3.

Точность формы (66) означает, что она является дифференциалом некоторой функции $S(t, x)$, определенной на G :

$$S_x(t, x) = p(t, x); \quad S_t(t, x) = -H(t, x, p(t, x)).$$

Тем самым, функция $S(t, x)$ является решением уравнения Гамильтона-Якоби. При этом $S|_N = \text{const}$ (интеграл от формы (66) по любой кривой $l \subset N$ равен нулю).

Итак, интеграл от формы (66) не зависит от пути γ в G . Перепишем этот интеграл в терминах подынтегральной функции f :

$$\begin{aligned} I(\gamma) &= \int_{\gamma} p(t, x)dx - H(t, x, p(t, x))dt = \\ &= \int_{\gamma} f_x(t, x, g(t, x))dx - [-f(t, x, g(t, x)) + f_x(t, x, g(t, x))g(t, x)]dt. \end{aligned} \quad (67)$$

Интеграл (67) называется инвариантным интегралом Гильберта для поля \mathcal{P} . Он не зависит от пути в области G , а на экстремальных поля, т. е. на решениях системы $\dot{x} = g(t, x)$, интеграл (67) совпадает с исходным функционалом:

$$\int_{\gamma} f_x dx - [-f + f_x g] dt = \int_{\gamma} f(t, x, g(x, t)) dt.$$

Функция Вейерштрасса для поля \mathcal{P} определяется на G формулой

$$\mathcal{E}_{\mathcal{P}}(t, x, \xi) = \mathcal{E}(t, x, g(t, x), \xi) = f(t, x, \xi) - f(t, x, g(t, x)) - p(t, x)(\xi - g(t, x)). \quad (68)$$

Теорема 7. (Достаточные условия сильного минимума в терминах поля экстремалей).

Пусть в области G определено поле экстремалей \mathcal{P} с начальным многообразием N , а функция Вейерштрасса (68) неотрицательна при всех $(t, x) \in G$ и при любых ξ . Пусть $\hat{x}(\cdot)$ – некоторая экстремаль поля \mathcal{P} . Тогда для любой кривой $x(\cdot)$ такой, что $x(t_1) = \hat{x}(t_1)$; $x(t_0) \in N$, $x(t) \in G$ при всех $t \in [t_0, t_1]$ имеет место неравенство $\mathcal{J}(\hat{x}(\cdot)) \leq \mathcal{J}(x(\cdot))$.

Доказательство.

В силу следствия 1: $I(\hat{x}(\cdot)) = I(x(\cdot))$. Но $\hat{x}(\cdot)$ – экстремаль и поэтому $I(\hat{x}(\cdot)) = \mathcal{J}(\hat{x}(\cdot))$. Следовательно, $\mathcal{J}(x(\cdot)) - \mathcal{J}(\hat{x}(\cdot)) = \mathcal{J}(x(\cdot)) - I(\hat{x}(\cdot)) = \mathcal{J}(x(\cdot)) - I(x(\cdot)) = \int_{t_0}^{t_1} f(t, x(t), \dot{x}(t)) dt - \int_{t_0}^{t_1} \{f_{\dot{x}}(t, x(t), g(t, x(t)))\dot{x}(t) - [-f(t, x(t), g(t, x(t))) + f_{\dot{x}}(t, x(t), g(t, x(t)))g(t, x(t))]\} dt = \int_{t_0}^{t_1} \mathcal{E}_{\mathcal{P}}(t, x(t), \dot{x}(t)) dt \geq 0$.

□

Итак, чтобы убедиться в том, что на экстремали $\hat{x}(\cdot)$ достигается сильный минимум, следует погрузить эту экстремаль в поле.

19 ЭЙЛЕРОВЫ ЭЛАСТИКИ

На плоскости (x, y) рассмотрим бесконечно тонкий, однородный, нерастяжимый, упругий стержень длины единица. Пусть этот стержень закреплен в концах так, что фиксированы направления, по которым он подходит к концам. Вопрос состоит в том, какую форму примет такой стержень. Примем за параметр длину дуги искомой кривой, а через ξ обозначим угол наклона единичного касательного вектора с осью абсцисс. В качестве управления $u = \dot{\xi}$ рассмотрим скорость изменения этого вектора, т.е. кривизну искомой кривой (модуль скорости единичного касательного вектора). Тогда $x(\xi), y(\xi)$ являются решениями управляемой системы

$$\begin{cases} \dot{x} = \cos \xi \\ \dot{y} = \sin \xi \\ \dot{\xi} = u \end{cases} \quad (69)$$

с граничными условиями

$$x(0) = x_0, y(0) = y_0, dy/dx(0) = p_0; \quad x(1) = x_1, y(1) = y_1, dy/dx(1) = p_1. \quad (70)$$

Упругая энергия стержня это интеграл от квадрата его кривизны

$$\frac{1}{2} \int_0^1 u^2 ds. \quad (71)$$

Изучим задачу минимизации интеграла (90) при условиях (88), (89). Функция Понтрягина для этой задачи имеет вид

$$H = \psi_1 \cos \xi + \psi_2 \sin \xi + \psi_3 u + \frac{\lambda}{2} u^2.$$

Сопряженная система имеет вид

$$\begin{cases} \dot{\psi}_1 = 0 \\ \dot{\psi}_2 = 0 \\ \dot{\psi}_3 = \psi_1 \sin \xi - \psi_2 \cos \xi \end{cases}$$

Множитель λ не равен нулю, т.к. иначе максимум функции H достигался бы только при $\psi_3 = 0$, но тогда, в силу последнего уравнения сопряженной системы, $\psi_1 = \psi_2 = 0$, и все множители Лагранжа были бы равны нулю. Поэтому можно положить $\lambda = -1$. Максимум H достигается при $u = -\psi_3$. В результате получается $\ddot{\xi} = -\rho \sin(\xi + \beta)$. Выбирая начало отсчета для угла ξ , можно считать, что $\beta = 0$. Итак, имеем уравнение физического маятника

$$\ddot{\xi} = -\rho \sin(\xi). \quad (72)$$

Умножая на $\dot{\xi}$ и интегрируя, получаем уравнение

$$\frac{1}{2}(\dot{\xi})^2 = \rho \cos(\xi) + C.$$

В зависимости от граничных условий решениями могут быть: прямая, когда при всех $i = 1, 2, 3$ имеем $\psi_i \equiv 0$; окружность, когда $\psi_1 = \psi_2 \equiv 0$; $\psi_3 \equiv 1$. При остальных значениях сопряженных переменных получается семейство эллиптических функций

$$dt = \frac{d\xi}{\sqrt{2C + 2\rho \cos \xi}} = \frac{d\xi}{\sqrt{(2C + 2\rho) - 4\rho \sin^2 \frac{\xi}{2}}}.$$

Интегрируя, получаем

$$\sqrt{\frac{2}{C + \rho}} \int_0^{\xi/2} \frac{d\eta}{\sqrt{1 - \frac{2\rho}{C + \rho} \sin^2 \eta}} = t + \gamma.$$

Выберем константу C так, чтобы $\frac{C + \rho}{2} = 1$. Тогда получаем

$$\xi = 2 \operatorname{am}((t + \gamma); k), \quad (73)$$

где $k^2 = \rho$. Модуль k будет фиксированным, и мы иногда будем его пропускать под знаком эллиптических функций. Проверим, что в результате действительно получилось решение уравнения (91)

$$\ddot{\xi} = -2\rho \operatorname{sn}(t + \gamma) \operatorname{cn}(t + \gamma) = -\rho \sin(2 \operatorname{am}(t + \gamma)).$$

Напомним уравнение Риккати для математического маятника [?]:

$$\begin{cases} \frac{d}{dt}q = p, \\ \frac{d}{dt}p = -q, \end{cases}$$

Уравнение в вариациях для линейного уравнения совпадает с ним самим, и для функции $w = -pq^{-1}$ получаем

$$\frac{dw}{dt} = 1 + w^2, \quad w = -\operatorname{tg}(t + c)$$

В этом простейшем случае пара взаимно сопряженных точек находится друг от друга на расстоянии π . Рассмотрим отдельно задачу для физического маятника (91). Его потенциальная энергия

равна $U = -\rho \cos \xi$, а полная энергия $H = \dot{\xi}^2/2 - \rho \cos \xi$. Следовательно, функция действия имеет вид $S = \dot{\xi}^2/2 + \rho \cos \xi$ и, в соответствии с принципом наименьшего действия, задача состоит в минимизации функционала

$$J = \int_0^T \left(\frac{\dot{\xi}^2}{2} + \rho \cos \xi \right) dt. \quad (74)$$

Уравнение Эйлера для этой задачи имеет вид (91). Вторая вариация функционала (74) на экстремали $\hat{\xi}(\cdot)$ имеет вид

$$\delta^2 J = \int_0^T \left(\frac{\dot{h}^2}{2} - \rho \cos \hat{\xi}(t) h^2 \right) dt. \quad (75)$$

Для изучения достаточных условий оптимальности применим метод уравнения Риккати. Прежде всего следует выписать уравнение Эйлера-Якоби, линеаризуя уравнение (91).

$$\ddot{h} = -(\rho \cos \xi)h.$$

Рассмотрим $W = -\dot{h}h^{-1}$. Продифференцировав, получаем уравнение Риккати

$$\dot{W} = -\ddot{h}h^{-1} + \dot{h}(h^{-1}\dot{h}h^{-1}) = \rho \cos(2 \operatorname{am}(t + \gamma)) + W^2.$$

Используя тригонометрическую формулу двойного угла и определение эллиптических функций, получаем

$$\dot{W} = \rho[\operatorname{cn}^2(t + \gamma) - \operatorname{sn}^2(t + \gamma)] + W^2. \quad (76)$$

Решения уравнения в вариациях можно получить, дифференцируя решения уравнения Эйлера по параметрам. Дифференцируя (92) по γ , получаем

$$h = 2 \operatorname{dn}(t + \gamma); \quad \dot{h} = -2\rho \operatorname{sn}(t + \gamma) \operatorname{cn}(t + \gamma).$$

Следовательно, решением уравнения Риккати (76) служит функция

$$W = -\dot{h}h^{-1} = \rho \frac{\operatorname{sn}(t + \gamma) \operatorname{cn}(t + \gamma)}{\operatorname{dn}(t + \gamma)}. \quad (77)$$

Уравнения Риккати получили свое название по имени знаменитого итальянского математика, графа Якопо Франческо Риккати.

Чтобы оценить степень его влияния на современников достаточно сказать, что многие университеты приглашали Риккати на должность профессора, Петр Великий предложил ему стать президентом Санкт-Петербургской Академии Наук, Венский Двор приглашал его на должность Имперского Канцлера. От всех этих предложений Риккати отказался, говоря, что предпочитает жить в кругу семьи и заниматься математикой. В 1724 г. Риккати опубликовал в журнале "Acta Eruditorum" статью "Animadversiones in aequationes differentiales secundi gradus посвященную методам разделения переменных и методам понижения порядка дифференциальных уравнений. Изучая задачу восстановления плоской кривой по свойствам ее кривизны, Риккати пришел в этой статье к рассмотрению уравнения

$$b \frac{dx}{dt} = at^\alpha + x^2. \quad (78)$$

Это уравнение допускает разделение переменных при

$$\alpha = \frac{4n}{-2n + 1},$$

где n — целое число. При $n \rightarrow \infty$ из этой формулы получается значение $\alpha = -2$, при котором уравнение (78) также интегрируется. В том же журнале появилась статья 22-летнего Даниила Бернулли, который писал, что оба Николая Бернулли (старший и младший), Иоганн Бернулли и он сам, Даниил Бернулли, изучали это уравнение и нашли все значения параметра α , при которых оно интегрируется. Риккати фактически отказался от спора о приоритете, заявив, что публикуя статью, он вовсе не имел намерения бросать вызов высоко почитаемому им семейству Бернулли. В 1841 г. Лиувилль доказал, что при всех остальных значениях параметра α уравнение (78) не интегрируется в квадратурах. Выше найдено еще одно важное уравнение Риккати (77) (связанное уже с эллиптическими функциями), которое интегрируется в квадратурах.

Теорема 8. . 1. Если на полуинтервале $(0, T]$ функция $\text{dn}(t)$ не обращается в ноль, то экстремаль $\hat{\xi}(t)$ оптимальна. 2. Если функция $\text{dn}(t)$ обращается в ноль в двух точках t_* и t^* полуинтервала $[0, T)$, то экстремаль $\hat{\xi}(t)$ не оптимальна.

Доказательство. В случае 1 прибавим под знаком интеграла (75) полную производную

$$\frac{d}{dt} \left(\frac{\rho \text{sn}(t) \text{cn}(t)}{\text{dn}(t)} h^2 \right).$$

Если функция $h(\cdot)$ удовлетворяет нулевым граничным условиям, то значение функционала не изменится. В силу выполнения уравнения Риккати под интегралом возникнет полный квадрат, что доказывает положительность второй вариации (75). В случае 2 рассмотрим функцию

$$\begin{cases} \tilde{h} = \text{dn}(t) & \text{при } t \in [t_*, t^*], \\ \tilde{h} = 0 & \text{при } t \notin [t_*, t^*]. \end{cases}$$

Если бы вторая вариация (75) была неотрицательна, то функция \tilde{h} наряду с функцией $h = 0$ давала бы минимум функционалу (75), не удовлетворяя необходимым условиям минимума в угловых точках. \square Вернемся к задаче об эйлеровых эластичках. Решение (77) уравнения Риккати (76) ух-

одит в бесконечность при $\text{dn}(t + \gamma) \rightarrow 0$. При этом ни $\text{sn}(t + \gamma)$ ни $\text{cn}(t + \gamma)$ в ноль не обращаются. Функция $\text{dn}(t + \gamma)$ имеет прямое отношение к задаче об эластичках.

Утверждение 2. Необходимым условием для точки перегиба эластички служит равенство $\text{dn}(t + \gamma) = 0$.

Доказательство. Используем уравнения (88)

$$\frac{dy}{dx} = \frac{\cos 2 \text{am}(t + \gamma)}{\sin 2 \text{am}(t + \gamma)} = \frac{\text{cn}^2(t + \gamma) - \text{sn}^2(t + \gamma)}{2 \text{sn}(t + \gamma) \text{cn}(t + \gamma)}.$$

Далее

$$\frac{d^2 y}{dx^2} = \frac{d}{dt} \left(\frac{dy}{dx} \right) \left(\frac{dx}{dt} \right)^{-1}.$$

Следовательно,

$$\frac{d^2 y}{dx^2} = \frac{-\sin^2(2 \text{am}(t + \gamma)) 2 \text{dn}(t + \gamma) - \cos^2(2 \text{am}(t + \gamma)) 2 \text{dn}(t + \gamma)}{\sin^3(2 \text{am}(t + \gamma))} = \frac{2 \text{dn}(t + \gamma)}{\sin^3(2 \text{am}(t + \gamma))}.$$

□ Доказанный факт вполне понятен, т.к. в точке перегиба направление вращения угла ξ меняется на противоположное. Поэтому $\xi = 2 \operatorname{dn}(t + \gamma)$ меняет знак, проходя через ноль. Найденное решение уравнения Риккати (76) и утверждение 1 объясняют, почему сопряженные точки играют столь важную роль в вопросах об устойчивости эластик. Макс Борн в 1906 году произвел серию физических экспериментов. Он рассмотрел вопрос о достаточных условиях оптимальности эластик и доказал, что при отсутствии перегибов на них сопряженных точек нет. Отсутствие сопряженных точек гарантирует устойчивость формы элаستيку при малых изменениях граничных условий. Если постепенно менять граничные условия, вращая концы стержня, то при возникновении пары сопряженных точек локальный минимум перестанет достигаться. Существуют сколь угодно малые изменения его формы, которые приводят к меньшим значениям минимизируемого функционала. Положение стержня станет неустойчивым. Реальный стержень начнет двигаться и в практических экспериментах он скачком перепрыгнет в положение, устойчивое для образовавшихся граничных условий. Устойчивые конфигурации могут оказаться глобально неоптимальными, но чтобы перейти в точку глобального минимума, требуются большие отклонения. Для изучения глобального минимума надо рассмотреть точки Максвелла (cut off), в которые приходит несколько экстремалей с одним и тем же значением функционала. Исследование сопряженных точек и точек Максвелла для плоских эластик, использующее группы симметрий решений уравнения Эйлера, было проведено в работе Сачкова [?]. С помощью прямого вычисления экспоненциального отображения для второй вариации, им было доказано, что сопряженная точка к исходной точке лежит между первой и третьей точкой перегиба.

История вариационной задачи об эластике началась в 1691 году, когда Якоб Бернулли рассмотрел задачу о форме, которую должен принять тонкий, упругий, однородный стержень, поставленный вертикально, вделанный в пол и наклоненный под воздействием груза, закрепленного на его верхнем конце. Он получил уравнение

$$\frac{dy}{dx} = \frac{x^2}{\sqrt{1-x^4}},$$

не интегрируемое в элементарных функциях. Я.Бернулли разложил решение в ряд и провел численные оценки.

В 1742 году Даниил Бернулли в письме к Эйлеру предложил ему найти решение более общей, плоской, вариационной задачи: Найти форму упругого стержня, минимизирующую функционал

$$\int \frac{ds}{R^2},$$

где R радиус кривизны искомой кривой. Эйлер в это время писал книгу по вариационному исчислению, опубликованную в 1744 году. Откликнувшись на призыв Бернулли, Эйлер получил дифференциальное уравнение для решения, интегрируемое только в эллиптических функциях. Исследование эллиптических функций было осуществлено Абелем и Якоби спустя почти полвека. Тем не менее, Эйлер описал и расклассифицировал все возможные формы, которые может принять стержень. С этого времени эти кривые получили название эйлеровых эластик.

Впоследствии и до самого последнего времени многие математики и физики возвращались к исследованию Эйлеровых эластик. Например, Гриффитс и Брайент, развивая Картановскую технику продолжения систем внешних дифференциальных форм применительно к задачам вариационного исчисления, использовали задачу об Эйлеровых эластике в качестве основной модельной задачи [?].

По таинственным причинам эйлеровы эластики возникают при решении самых разнообразных физических задач, казалось бы, совершенно не связанных с задачей о минимизации интеграла от

квадрата кривизны (??). В этой простой задаче поразительным образом сплелись вопросы теории упругости, гидро и газовой динамики, магнетизма, квантовой механики и т.п.. Перечислю лишь некоторые примеры.

В 1986 году Артур и Волш (а также позднее и независимо Джорджевич), исследуя задачу о качении шара по плоскости без проскальзывания, обнаружили следующий феномен. При решении задачи о минимизации длины кривой контакта с плоскостью при оптимальном переводе шара из одного пространственного положения в другое, оптимальными кривыми контакта оказываются эйлеровы эластики [?].

Кирхгоф в середине XIX века рассмотрел динамические задачи о вихревых течениях невязкой, несжимаемой жидкости (или газа). Он исследовал течения, имеющие форму вихревых трубок, которые он назвал *vortex filament* (вихревые нити). Кирхгоф показал, что их уравнения определяются эллиптическими функциями такими же, как эйлеровы эластики. Поэтому в ряде современных работ они называются кирхгофовыми (а не эйлеровыми) эластиком. В частности, оси вращения торнадо в атмосфере описываются уравнениями, связанными с эйлеровыми эластиком.

В 1906 году флорентийский математик и механик Да Риос, ученик Леви-Чивита, исследовал движение вихревых нитей и получил замечательное уравнение, описывающее их эволюцию [?]. Пусть функция $\gamma(s, t)$ в фиксированный момент времени t это уравнение вихревой кривой, отнесенной к натуральному параметру s , т.е. к длине дуги. Уравнение Да Риоса имеет вид

$$\gamma_t = \gamma_s \wedge \gamma_{ss}. \quad (79)$$

Здесь нижний индекс означает производную по соответствующему аргументу, знак внешнего умножения соответствует в трехмерном случае векторному произведению, γ_s единичный касательный вектор к вихревой кривой, а γ_{ss} вектор, направленный по главной нормали, длина которого равна модулю кривизны. Значит на ось вихря действует усилие пропорциональное кривизне и направленное по бинормали. Причиной этой эволюции является взаимодействие течений (в соответствии с законом Бернулли) по ближайшим наклоненным друг к другу кругам, ортогональным оси вихря. Поэтому уравнение (79) часто называют LIA (Local Induction Approximation) или потоком бинормалей. Его связывают с одним из видов турбулентности, так называемой вихревой турбулентностью.

Замечание 4. *С точки зрения изучения движения тайфуна было бы интересно найти правильные граничные условия и получить, используя формулу (??), законы сохранения для незамкнутых кривых. Хорошо было бы их сравнить с законами сохранения для нелинейного уравнения Шредингера (см. ниже о преобразовании Хасимото).*

Уравнение Да Риоса возникает при изучении эволюции одномерной спин системы (фермионной жидкости, описывающей магнитное взаимодействие фермионов, обладающих ненулевым спином) [?].

Особенно существенным уравнение Да Риоса становится для плазмы, когда в результате круговых токов вокруг оси вихря возникает магнитное поле направленное вдоль оси "катушки". В работах по астрофизике высказывалось мнение, что солнечные протуберанцы это вихревые нити в плазме солнечной короны. Видимо, по этой причине вихревые нити стали интенсивно изучаться в середине XX-го века в теории магнетизма. Было выяснено [?], что силовые линии так называемых киллинговых магнитных полей (т.е. для полей, сдвиги по траекториям которых определяют изометрии пространства) удовлетворяют уравнению (79). Релятивистские заряженные частицы под влиянием собственного магнитного момента и киллинговых магнитных полей на однородных

многообразиях постоянной кривизны (т.е. в ситуациях, описываемых общей теорией относительности) движутся по траекториям эйлеровых эластик. Поэтому естественным является появление эйлеровых эластик в теории сверхпроводимости.

Как известно, при воздействии на сверхпроводник достаточно сильного магнитного поля происходит фазовый переход в нормальное (не сверхпроводящее) состояние. Сверхпроводимость исчезает не мгновенно. При магнитном поле близком к критическому сначала возникает смешанное состояние. В основной своей массе проводник остается в сверхпроводящем состоянии, но в нем возникают, так называемые, абрикосовские нити. Это вихревые линии, центральная область которых, подобно глазу тайфуна, перешла из сверхпроводящего в нормальное состояние. Эти нити взаимодействуют друг с другом. Они отталкиваются. В проводнике возникает аналог кристаллической решетки, состоящей не из атомов, а из нитей. Эти нити имеют форму эластик. В физических экспериментах на тонких пленках эту структуру можно явно наблюдать, так что это не только теоретические конструкции. Правда, надо сказать, что мне не известны эксперименты, в которых можно было бы наблюдать форму самих нитей. В экспериментах фиксируются только центры — следы нитей на тонкой пленке, образующие красивый кристаллический узор. При дальнейшем повышении напряженности внешнего магнитного поля вихревые нити сливаются, и проводник целиком переходит в нормальное состояние.

В 80-ых годах двадцатого века японский математик Хасимото предложил преобразование, которое переводит решение вихревого уравнения Да Риоса (79) в решение знаменитого, вполне интегрируемого нелинейного уравнения Шредингера, играющего важную роль в квантовой оптике

$$iu_t = -u_{xx} - \frac{1}{2}u|u|^2.$$

Это уравнение обладает полным, счетным набором первых интегралов и имеет солитонные решения.

Преобразование Хасимото отождествляет пространство связных, гладких кривых $\gamma(s)$ без точек распрямления, профакторизованного по группе жестких движений пространства \mathbb{R}^3 , с множеством комплексных функций ψ . Отождествление задается комплекснозначной функцией $\psi(s)$, называемой в квантовой механике волновой функцией, определяющей состояние системы. Волновая функция задается формулой

$$\psi(s) = k(s) \exp\left(i \int_0^s \kappa(\xi) d\xi\right),$$

где $k(s)$ кривизна, а $\kappa(s)$ кручение кривой γ . Обратное отображение определяется формулами

$$k(s) = |\psi(s)|, \quad \kappa(s) = \frac{d}{ds} \arg \psi(s).$$

Мы видим, насколько разнообразные применения имеет решение этой несложной, но столь важной, экстремальной задачи.

20 МИНИМАЛЬНЫЕ ПОВЕРХНОСТИ

Важную роль в теории минимизации кратного интеграла сыграли работы Плато по минимальным поверхностям. Кроме того, к задаче минимизации кратного интеграла приводят вариационные принципы для квантовой теории поля, для сплошной среды (теории упругости и пластичности, теория оболочек), и т.д..

Граница подмногообразий для простоты предполагается фиксированной.

Пусть V – область n -мерного евклидова пространства с краем ∂V ; X – есть ν -мерное линейное пространство. Координаты на V будут обозначаться через $t = (t^1, \dots, t^n)$, координаты на X – через $x = (x^1, \dots, x^\nu)$.

Рассмотрим интегральный функционал

$$\mathcal{J} = \int_V f \left(t, x, \frac{\partial x}{\partial t} \right) dt, \quad (80)$$

где dt есть элемент объема многообразия V : $dt = dt^1 \wedge \dots \wedge dt^n$. Пусть $f : (\mathbb{R} \times \mathbb{R}^\nu \times M^{n\nu}) \rightarrow \mathbb{R}$ – гладкая функция (класса C^∞) в окрестности графика многообразия $\{t \in V, x = \hat{x}(t), u = \frac{\partial \hat{x}}{\partial t}(t)\}$. Здесь $M^{n\nu}$ – множество матриц размерности $n \times \nu$. В дальнейшем используется соглашение о суммировании по повторяющимся индексам, причем латинские индексы пробегают значения от 1 до n ; греческие индексы пробегают значения от 1 до ν . Функционал \mathcal{J} рассматривается на пространстве $C^1(V, \mathbb{R}^\nu)$ – непрерывно дифференцируемых функций $x(\cdot)$ на множестве V , удовлетворяющих граничным условиям

$$x \Big|_{\partial V} = \phi(t), \quad (81)$$

где $\phi(t)$ есть заданная гладкая функция.

Теорема 9. Пусть $x(\cdot) \in C^2(V, \mathbb{R}^\nu)$ есть функция, в которой достигается слабый минимум функционала (80), (81). Тогда $x(\cdot)$ удовлетворяет уравнению Эйлера

$$-\frac{\partial}{\partial t^\alpha} \frac{\partial f}{\partial \left(\frac{\partial x^i}{\partial t^\alpha} \right)} + \frac{\partial f}{\partial x^i} = 0. \quad (82)$$

Докажем сначала лемму об интегрировании по частям. Обозначим через $\{dt\}_\beta$ внешнее произведение всех дифференциалов dt^α (при естественном порядке сомножителей), в котором пропущен сомножитель dt^β .

Лемма 3. (Об интегрировании по частям).

Пусть $a(t), b(t)$ – гладкие функции на области V . Имеет место формула

$$\int_V \left[a \frac{\partial b}{\partial t^\beta} \right] dt = - \int_V \left[b \frac{\partial a}{\partial t^\beta} \right] dt + \int_{\partial V} (-1)^\beta ab \{dt\}_\beta. \quad (83)$$

Доказательство.

Рассмотрим дифференциальную форму $\omega = a(t)b(t)\{dt\}_\beta$, определенную на области V . Применим к ней формулу Стокса

$$\int_V d\omega = \int_{\partial V} \omega.$$

Имеем

$$\int_V a \frac{\partial b}{\partial t^\beta} dt + \int_V b \frac{\partial a}{\partial t^\beta} dt = (-1)^\beta \int_{\partial V} ab \{dt\}_\beta.$$

Лемма доказана.

□

Доказательство теоремы. Пусть в точке $x(\cdot)$ достигается слабый минимум функционала (80). Продифференцируем в этой точке функционал (80) по направлению $h(\cdot)$, удовлетворяющему условию

$$h|_{\partial V} = 0. \quad (84)$$

Вариации $x(\cdot) + \lambda h(\cdot)$, где $\lambda \in R^1$, удовлетворяют условию (81) и при достаточно малом λ попадают в любую наперед заданную C^1 -окрестность функции $x(\cdot)$. Подставив эти вариации в функционал (80), получаем скалярную функцию $\mathcal{J}(\lambda)$, которая при $\lambda = 0$ достигает минимального значения, и, следовательно, $\mathcal{J}'(0) = 0$. Имеем

$$\mathcal{J}'(0) = \int_V \left(\frac{\partial f}{\partial x^i} h^i + \frac{\partial f}{\partial \left(\frac{\partial x^i}{\partial t^\alpha} \right)} \frac{\partial h^i}{\partial t^\alpha} \right) dt. \quad (85)$$

Проинтегрируем второе слагаемое подинтегрального выражения в формуле (85) по частям. Получим

$$\int_V \left(-\frac{\partial}{\partial t^\alpha} \left(\frac{\partial f}{\partial \left(\frac{\partial x^i}{\partial t^\alpha} \right)} \right) + \frac{\partial f}{\partial x^i} \right) h(t) dt = 0$$

для любой функции $h(\cdot)$, удовлетворяющей условию (84). Отсюда немедленно следует, что коэффициент при h тождественно равен нулю, что и требовалось доказать.

□

Задача Плато о поверхности минимальной площади в трехмерном евклидовом пространстве, натянутой на заданный контур, формулируется следующим образом

Минимизировать функционал

$$\int_{\Omega} \sqrt{1 + (z_x)^2 + (z_y)^2} dx dy \quad (86)$$

по множеству гладких функций $z = z(x, y)$, определенных на области $\Omega \subset R^2$ и удовлетворяющих граничным условиям

$$h|_{\partial\Omega} = \phi(t).$$

Применяя уравнение Эйлера (82) к функционалу (86), получаем после несложных преобразований уравнение

$$z_{xx}(1 + (z_y)^2) - 2z_{xy}z_xz_y + z_{yy}(1 + (z_x)^2) = 0. \quad (87)$$

Упражнение 1. Проверить, что катеноид (поверхность вращения кривой $y = C \operatorname{sch}\left(\frac{x}{c}\right)$ вокруг оси абсцисс) удовлетворяет уравнению (87).

Упражнение 2. (Геликоид.)

Геликоид получается в результате композиции двух движений прямой линии: поступательного с постоянной скоростью и вращательного с постоянной угловой скоростью в плоскости,

ортогональной вектору скорости поступательного движения. Иными словами, z пропорционально углу поворота прямой. Следовательно, уравнения геликоида (с точностью до множителя) имеют вид

$$z = \operatorname{arctg} \frac{y}{x} .$$

Проверить, что геликоид дает решение уравнения (87).

Поверхность Шерка.

Найдем решение уравнения (87), записывающееся в виде $z = f(x) + g(y)$. Подставив это выражение в (87), получим

$$f''(x)(1 + g'(y)^2) + g''(y)(1 + f'(x)^2) = 0.$$

Разделяя переменные, получаем

$$\frac{f''(x)}{(1 + f'(x)^2)} = c, \quad \frac{g''(y)}{(1 + g'(y)^2)} = -c.$$

Интегрирование этих уравнений (с точностью до сдвигов по осям координат и до гомотетии) дает

$$z = \ln \frac{\cos y}{\cos x} .$$

21 ПРАВИЛО МНОЖИТЕЛЕЙ ЛАГРАНЖА ДЛЯ ГЛАДКИХ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ

Для доказательства правила множителей Лагранжа для гладких экстремальных задач нам потребуется серия вспомогательных лемм и теорем.

Теорема отделимости.

Теорема 10. Пусть A и B два выпуклых множества в банаховом пространстве X . Тогда, если

1. $A \cup B = \emptyset$,
2. Хотя бы в одном из множеств A или B есть внутренняя точка.

Тогда множества A и B отделимы.

Два множества $A \in X$ и $B \in X$ называются отделимыми, если существует разделяющая их гиперплоскость, т.е. элемент $W \in X^*$ такой, что $\inf_{x \in A} (W, x) \geq \sup_{y \in B} (W, y)$.

Доказательство см. [Колмогоров-Фомин с.137]. Заметим, что условие 2. требуется только для бесконечномерных пространств X . В конечномерном случае теорема верна и без этого условия.

Поясним это на примере.

Рассмотрим компактный эллипсоид $A = \sum n^2 x_n^2 \leq 1$ в пространстве l_2 .

Задача 1. Доказать, что A компакт.

Задача 2. При каких a_n эллипсоид $\sum \frac{x_n^2}{a_n^2} \leq 1$ является компактом.

Рассмотрим луч $B = t\{\frac{1}{n}\}$, где $t > 0$. Проверим для A и B выполнение условий теоремы отделимости. Луч, очевидно, выпуклое множество. Выпуклость эллипсоида проверяется следующей выкладкой. Пусть $x = \{x_n\}$ и $y = \{y_n\}$ две точки эллипсоида. Берем $z = \{\alpha x_n + \beta y_n\}$, где $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$. Имеем

$$\begin{aligned} \sum n^2 (\alpha x_n + \beta y_n)^2 &= \alpha^2 \sum n^2 x_n^2 + 2\alpha\beta \sum n^2 x_n y_n + \beta^2 \sum n^2 y_n^2 \leq \\ &\alpha^2 + \beta^2 + \alpha\beta (\sum n^2 x_n^2 + \sum n^2 y_n^2) \leq (\alpha + \beta)^2 \leq 1. \end{aligned}$$

Далее, ряд $\sum n^2 \frac{t}{n^2}$ расходится, и, следовательно, луч B не пересекается с границей эллипсоида. Мы имеем чисто бесконечномерный эффект: несмотря на то, что поверхность эллипсоида, казалось бы, окружает начало координат, можно дойти до этого начала по прямой, не задевая поверхности. Следовательно, $A \cup B = \emptyset$. Покажем, что A и B неотделимы. Действительно, если бы существовал разделяющий функционал $W = \{W_n\} \in l_2^* = l_2$, то $\inf_{x \in B} (W, x) = 0$, но тогда и $\sup_{y \in A} (W, y) = 0$ и весь эллипсоид лежал бы в полупространстве $(W, x) > 0$. Но A центрально симметричен. Значит он должен целиком принадлежать подпространству W . В принципе, "нормаль" к этому подпространству – вектор $\{w_n\}$ – как мы видели, может и не пересекать границу эллипсоида, но, по крайней мере, вектор $\{\frac{w_n}{n}\}$ пересекается с границей, так как соответствующий ряд $\sum n^2 \frac{(tw_n)^2}{n^2}$ сходится. Но это пересечение не лежит в плоскости $(W, x) = 0$. Противоречие.

Причина в том, что ни луч ни эллипсоид не содержат внутренних точек. для эллипсоида отсутствие внутренних точек связано с тем, что у него есть бесконечно малые полуоси, и он пересекается с границей любого шара в l_2 с центром во внутренней точке эллипсоида.

ЛЕММА О НЕТРИВИАЛЬНОСТИ АННУЛЯТОРА.

Определение 9. Подпространством банахова пространства X называется замкнутое линейное подмножество X .

Определение 10. Аннулятором подпространства $L \in X$ называется множество L^\perp линейных функционалов $x^* \in X^*$, таких что $(x^*, x) = 0$ при любых $x \in L$.

Лемма 4. $\{L \neq X\} \Rightarrow \{L^\perp \neq \emptyset\}$

Доказательство.

Пусть $a \notin L$. Поскольку L замкнуто, то существует O – окрестность точки a , которая не пересекается с L . Следовательно, по теореме отделимости O и L разделимы, т.е. существует такой элемент $x^* \neq 0$, что $(x^*, l) < (x^*, a)$ при любом $l \in L$. Тогда $(x^*, l) = 0$, потому что в противном случае $(x^*, \lambda l)$ при различных знаках скаляра λ менялось бы от $+\infty$ до $-\infty$.

□

Лемма о правом обратном.

Лемма 5. Пусть $A : X \rightarrow Y$ сюръективный непрерывный линейный оператор из банахова пространства X в банахово пространство Y .

Тогда существует отображение (не обязательно линейное) $M : Y \rightarrow X$ и константа $C > 0$ такие, что $A \circ M = Id$, причем $\|My\| \leq C \|y\|$.

Доказательство.

Поскольку A непрерывен, то $Ker A$ подпространство. Рассмотрим факторпространство $X/Ker A$ и оператор $\tilde{A} : X/Ker A \rightarrow Y$, определенный следующим образом $\tilde{A}(\xi) = Ax$, где $x \in \xi$. Легко видеть, что это определение корректно, т.к. $\tilde{A}(\xi)$ не зависит от выбора элемента $x \in \xi$. Оператор \tilde{A} это непрерывный, линейный, сюръективный оператор из банахова пространства $X/Ker A$, действующий в Y инъективно. По теореме Банаха у него существует непрерывный обратный \tilde{A}^{-1} , и $\|\tilde{A}^{-1}\| = K$. Определим $M : Y \rightarrow X$, выбрав из смежного класса $\tilde{A}^{-1}(y)$ элемент $x(y)$ так, чтобы $\|x(y)\| \leq 2 \|\tilde{A}^{-1}(y)\|$. Ясно, что M правый обратный, т.к. $Ax(y) = y$ для любого элемента x из класса смежности $\tilde{A}^{-1}(y)$ и $\|My\| = \|x(y)\| \leq 2 \|\tilde{A}^{-1}(y)\| \leq 2K \|y\|$.

□

ЛЕММА О ЗАМКНУТОСТИ ОБРАЗА

Лемма 6. Пусть A и B линейные непрерывные операторы, действующие из банахова пространства X в банаховы пространства Y и Z соответственно. Пусть образ оператора A — ImA — замкнут и образ ядра A при действии оператора B переходит в замкнутое множество.

Тогда оператор $C : X \rightarrow Y \times Z$ (где $Cx = (Ax, Bx)$) имеет замкнутый образ.

Доказательство.

Пусть $(y_n, z_n) \in ImC$ и $(y_n, z_n) \rightarrow (y_0, z_0)$. Имеем $Ax_n = y_n$, $Bx_n = z_n$. Надо показать, что существует x_0 , такой, что $Ax_0 = y_0$, $Bx_0 = z_0$. Взяв вместо Y банахово пространство ImA , мы можем без ограничения общности считать, что A сюръективен. Следовательно он имеет правый обратный оператор $M : A \circ M = Id$, $\|My\| \leq K \|y\|$. Подправим последовательность x_n на бесконечно малую величину так, чтобы она целиком лежала в прообразе точки y_0 . Для этого берем $\tilde{x}_n = x_n - M(y_n - y_0)$. Тогда $A\tilde{x}_n = Ax_n - y_n + y_0 = y_0$, при этом $\|x_n - \tilde{x}_n\| = \|y_n - y_0\| \rightarrow 0$. Поэтому $B(x_n - \tilde{x}_n) \rightarrow 0$ и $B(\tilde{x}_n) \rightarrow z_0$. Все точки \tilde{x}_n лежат в аффинном подмножестве Ξ параллельном ядру A , т.к. они имеют общий A -образ. Следовательно, образ Ξ при действии оператора B замкнут. Коль скоро $B(\tilde{x}_n) \rightarrow z_0$, то существует $x_0 \in \Xi$, такой что $Bx_0 = z_0$. Итак, $Ax_0 = y_0$, $Bx_0 = z_0$.

□

ЛЕММА ОБ АННУЛЯТОРЕ ЯДРА

Лемма 7. Пусть A линейный, непрерывный, сюръективный оператор из банахова пространства X в банахово пространство Y .

Тогда $(KerA)^\perp = ImA^*$.

Доказательство.

1. Пусть $x^* \in ImA^*$, т.е. $x^* = A^*y^*$. Тогда для $h \in KerA$ имеем $(x^*, h) = (A^*y^*, h) = (y^*, Ah) = 0$.

2. Пусть $h \in (KerA)^\perp$. Рассмотрим оператор $C : X \mapsto \{Ax, (h, x)\}$. Поскольку $h(KerA) = 0$, то по лемме о замкнутости образа, оператор C имеет замкнутый образ. ImC не совпадает со всем пространством $\{y, \mathbb{R}\}$, т.к. $\{0, 1\} \notin ImC$. По лемме о нетривиальности аннулятора $C^\perp \neq \emptyset$. Поэтому существуют ненулевые $\{y^*, \lambda\}$, такие что $(y^*, Ax) + \lambda(h, x) = 0$ для любого x . Заметим, что $\lambda \neq 0$, т.к. иначе ненулевой y^* аннулировал бы все X . Следовательно, $((\frac{1}{\lambda}A^*y^* + h), x) = 0$, т.е. $h \in ImA^*$.

□

ТЕОРЕМА ЛЮСТЕРНИКА

Теорема 11. Пусть $F : X \rightarrow Y$ отображение класса C^1 , заданное в окрестности точки x_0 . Пусть $F'(x_0)$ сюръективно. Тогда существует такая функция $\varphi : X \rightarrow X$, что $F(x + \varphi(x)) = F(x_0)$, причем $\|\varphi(x)\| \leq C \|F(x) - F(x_0)\|$

Перенося начало координат в пространствах X и Y , можно без ограничения общности считать, что $x_0 = 0$ и $F(x_0) = 0$. Итак, надо чтобы $F(x + \varphi(x)) = 0$ и $\|\varphi(x)\| \leq C \|F(x)\|$.

Будем использовать метод Ньютона для поиска неподвижных точек отображения. Простейший вариант метода Ньютона состоит в использовании рекуррентной формулы $y_n = y_{n-1} - \frac{F(y_{n-1})}{F'(0)}$. Используя лемму о правом обратном, рассмотрим обобщение этой формулы в виде

$$\xi_n - \xi_{n-1} = MF(\xi_{n-1}), \quad (88)$$

где M правый обратный к оператору $F'(0)$. Для того чтобы доказать, что эта формула определяет сжимающее отображение, применим к обеим частям оператор $F'(0)$:

$$F(\xi_{n-1}) + F'(0)(\xi_n - \xi_{n-1}) = 0 \quad (89)$$

и оценим левую часть уравнения (88)

$$\| \xi_{n+1} - \xi_n \| \leq C \| F(\xi_n) \| . \quad (90)$$

Преобразуем правую часть уравнения (90), добавив под знаком нормы группу слагаемых, равную нулю

$$C \| F(\xi_n) \| = C \| F(\xi_n) - F(\xi_{n-1}) - F'(0)(\xi_n - \xi_{n-1}) \| . \quad (91)$$

По теореме о среднем правая часть (91) оценивается выражением

$$C \| F'(\zeta) - F'(0) \| \| \xi_n - \xi_{n-1} \| ,$$

где $\zeta \in [\xi_{n-1}, \xi_n]$. Поскольку $F \in C^1$, найдется шар B_r радиуса r такой, что $\|F'(\zeta) - F'(0)\| \leq 1/2C$ при $\zeta \in B_r$. Тогда, если $\xi_{n-1}, \xi_n \in B_r$, то

$$\| \xi_{n+1} - \xi_n \| \leq \frac{1}{2} \| \xi_n - \xi_{n-1} \| . \quad (92)$$

Из неравенства (92) следует сходимость ξ_n , если будет установлено, что для всех достаточно малых $x = \xi_0$ все $\xi_n \in B_r$. Выбираем для этого окрестность U точки x_0 такую, что $\|x\| + C \|F(x)\| < r/2$. Оцениваем последовательно $\| \xi_0 \|$, $\| \xi_0 \| + \| F(\xi_0) \| < r/2, \Rightarrow \| \xi_1 \| < r/2, \| \xi_1 - \xi_0 \| < C \| F(\xi_1) \| < r/2$. Далее $\| \xi_2 \| \leq \frac{1}{2} \| \xi_1 - \xi_0 \| < r/4$ и по индукции: Пусть при $i < N$ $\| \xi_i \| < r$. Тогда $\| \xi_N \| = \| \xi_N - \xi_{N-1} \| + \| \xi_{N-1} - \xi_{N-2} \| \dots + \| \xi_1 \| < r/2 + r/4 + \dots < r$. Точно так же доказывается фундаментальность, т.е. $\| \xi_{n+p} - \xi_n \| < r/2^{n-1}$ независимо от p . Итак, имеем сходимость и предел $\lim_{n \rightarrow \infty} \xi_n = \xi(x)$ зависит от выбора начальной точки $\xi_0 = x$. Введем $\varphi(x) = \xi(x) - x$. Переходя к пределу в формуле $F(\xi_n) = -F'(0)(\xi_n - \xi_{n-1}) \rightarrow 0$, получаем $F(\xi(x)) = F(x + \varphi(x)) = 0$. Переходя к пределу в неравенстве $\| \xi_n - x \| \leq 2 \| \xi_1 - x \|$, получаем $\| \xi(x) - x \| = \| \varphi(x) \| \leq 2C \| F(x) \|$.

□

ТЕОРЕМА О КАСАТЕЛЬНОМ ПРОСТРАНСТВЕ

Определение 11. Вектор h называется касательным к множеству M в точке x_0 , если существует функция $r(t) = o(t)$ такая что $x_0 + ht + r(t) \in M$.

Для касательных векторов используются обозначения $h \in T_{x_0}M$.

Пусть $F : X \rightarrow Y$ линейный, непрерывно дифференцируемый оператор из банахова пространства X в банахово пространство Y . Обозначим через M нулевую поверхность уровня F . Пусть в точке $x_0 \in M$ оператор $F'(x_0)$ сюръективен.

Теорема 12. Пусть $F : X \rightarrow Y$ отображение класса C^1 , заданное в окрестности точки x_0 . Пусть $F'(x_0)$ сюръективно. Тогда $\text{Ker}F'(x_0) = T_{x_0}M$.

Доказательство.

1. Пусть $h \in \text{Ker}F'(x_0)$. Используя функцию φ из теоремы Люстерника, имеем $F(x_0 + ht + \varphi(x_0 + ht)) = 0$, и $\| \varphi(x) \| \leq C \| F(x_0) + ht \|$. Но $\varphi(x) = F(x_0) + F'(x_0)ht + o(t) = o(t)$. Следовательно, $h \in T_{x_0}M$.

2. Пусть $h \in T_{x_0}M$. Тогда существует $r(t) = o(t)$ такое, что $F(x_0) + ht + r(t) = 0$. Учитывая, что $F(x_0) = 0$, разделив на t и перейдя к пределу при $t \rightarrow 0$, получим, что $h \in \text{Ker}F'(x_0)$.

□

ВЫВОД ПРАВИЛА МНОЖИТЕЛЕЙ ЛАГРАНЖА

Пусть на банаховом пространстве X заданы гладкая скалярная функция $f : X \rightarrow \mathbb{R}$ и гладкий оператор $F : X \rightarrow Y$.

Задача 3. Минимизировать функцию f на множестве $M = \{F(x) = 0\}$.

Теорема 13. Пусть x_0 решение задачи 1. Тогда, если $ImF'(x_0)$ замкнутое подпространство пространства X , то существует нетривиальный набор множителей Лагранжа $\lambda \in \mathbb{R}, y^* \in Y^*$ такой, что $\lambda f'(x_0) + (y^*, F'(x_0)) = 0$.

Доказательство.

1. Нерегулярный случай.

Пусть $ImF'(x_0) \neq Y$. Тогда, по лемме о нетривиальности аннулятора существует ненулевой функционал y^* , аннулирующий Y . следовательно, набор множителей Лагранжа $\{0, y^*\}$ удовлетворяет условиям теоремы.

2. Регулярный случай $ImF'(x_0) = Y$. Тогда $F'(x_0)$ сюръективно и по теореме об аннуляторе ядра $KerF'(x_0) = T_{x_0}M$. Покажем, что $f' \in KerF'$. При любом $h \in KerF'(x_0)$ имеем $h \in T_{x_0}M$, следовательно, существует $r(t) = o(t)$ такое, что $F(x_0 + ht + r(t)) = 0$. Но функция $f|_M$ достигает минимума в точке x_0 , т.е. при $t = 0$, значит $f'(x_0)h = 0$, т.е. $f'(x_0) \in (KerF'(x_0))^\perp$ или $f' \in Im(F')^*$. Значит существует элемент y^* такой, что $f' = (F')^*y^*$, т.е. набор $\{1, -y^*\}$ удовлетворяет условиям теоремы.

□

22 ПРИНЦИП МАКСИМУМА ПОНТРЯГИНА

ФОРМУЛИРОВКА ПРИНЦИПА МАКСИМУМА

Задача 4. $B_0(\xi) \rightarrow \min$

$$B_i(\xi) \leq 0, \quad (i = 1 \dots k)$$

$$B_j(\xi) = 0, \quad (j = k + 1 \dots m)$$

$$\dot{x}(t) - \varphi(t, x(t), u(t)) = 0$$

$$u(t) \in U$$

где $\xi = (x(\cdot), u(\cdot), t_0, t_1)$, функция $u(\cdot)$ измеримая, функция $x(\cdot)$ абсолютно непрерывная и

$$B_i(\xi) = \int_{t_0}^{t_1} f_i(t, x(t), u(t)) dt + l_i(t_0, x(t_0), t_1, x(t_1))$$

Определение 12. Говорят, что управляемый процесс $\hat{\xi}$ доставляет сильный локальный минимум, если существует такое $\delta > 0$, что $B_0(\xi) \geq B_0(\hat{\xi})$ для любого допустимого процесса $\xi = (x(\cdot), u(\cdot), t_0, t_1)$, для которого

$$|t - \hat{t}_0| < \delta, |t - \hat{t}_1| < \delta, \|x(t) - \hat{x}(t)\| < \delta.$$

Теорема 14. Пусть все рассматриваемые функции непрерывно дифференцируемы и процесс $\hat{\xi} = (\hat{x}(\cdot), \hat{u}(\cdot), \hat{t}_0, \hat{t}_1)$ оптимален. Тогда существуют такие множители Лагранжа $(\lambda, p(t))$, что для функции Лагранжа

$$\Lambda = \int_{t_0}^{t_1} (f(t, x, u) + p(t)(\dot{x} - \varphi(t, x, u))) dt + l(t_0, x(t_0), t_1, x(t_1))$$

выполнены следующие условия

- а). Уравнение Эйлера для лагранжиана $L = f(t, x, u) + p(\dot{x} - \varphi(t, x, u))$:
 $-\frac{d}{dt}L_{\dot{x}} + L_x = 0$;
 б). Условие трансверсальности: $L_{\dot{x}}(t_0) = l_{x(t_0)}$, $L_{\dot{x}}(t_1) = -l_{x(t_1)}$;
 в). Условие оптимальности: $\min_{u \in U} L(t, \hat{x}(t), u) = L(t, \hat{x}(t), \hat{u}(t))$;
 г). Условия дополняющей нежесткости и неотрицательности для $j \leq k$:
 $\lambda_j B_j(\hat{\xi}) = 0$, $\lambda_j \geq 0$.

Замечание 5. С точки зрения исторической традиции и в соответствии с самим названием "принцип максимума мы будем считать $\lambda_0 \leq 0$ и заменим условие минимума в пункте в). на условие максимума.

23 ДОКАЗАТЕЛЬСТВО ПРИНЦИПА МАКСИМУМА ДЛЯ ЗАДАЧИ СО СВОБОДНЫМ КОНЦОМ

Рассмотрим задачу

Задача 5. $\int_{t_0}^{t_1} f(t, x(t), u(t))dt \rightarrow \min$
при условиях

$$\dot{x}(t) - \varphi(t, x(t), u(t)) = 0, \quad x(t_0) = x_0; \quad u(t) \in U. \quad (93)$$

Лагранжиан имеет вид

$$L = \int_{t_0}^{t_1} \lambda_0 f(t, x, u) + p(\dot{x} - \varphi(t, x, u))dt + \mu(x(t_0) - x_0).$$

Применяя принцип максимума, получаем

$$\dot{p} = \lambda_0 f_x - p\varphi_x; \quad p(t_0) = \mu, p(t_1) = 0; \quad \min_u [\lambda_0 f - p\varphi] = [\lambda_0 f(t, \hat{x}, \hat{u}) - p\varphi(t, \hat{x}, \hat{u})].$$

Легко видеть, что

$$\lambda_0 = 0 \Rightarrow p \equiv 0 \Rightarrow \mu = 0 \Rightarrow \lambda_0 = -1.$$

ЛЕММА О СОПРЯЖЕННЫХ СИСТЕМАХ

Лемма 8. Пусть

$$\begin{cases} \dot{z} = A(t)z \\ \dot{p} = -pA(t) + g(t). \end{cases}$$

Тогда

$$pAz(t_1) - pAz(t_0) = \int_{t_0}^{t_1} g(t)z(t)dt \quad (94)$$

Для доказательства достаточно найти производную

$$\frac{d}{dt}(pAz) = -pAz + pAz + gz.$$

Интегрируя от t_0 до t_1 , получаем (94)

□

ЛЕММА ОБ ИГОЛКЕ

Лемма 9. Пусть $\hat{x}(t)$, $\hat{u}(t)$ оптимальный процесс, τ точка непрерывности управления $u(t)$. Рассмотрим вариацию

$$u(t) = \begin{cases} \hat{u}(t) & \text{при } t \notin [\tau, \tau + \sigma] \\ v & \text{при } t \in [\tau, \tau + \sigma] \end{cases}$$

Траекторию, соответствующую этому управлению, обозначим через $x(t, \sigma)$.

Тогда

1. $x(t, \sigma)$ равномерно стремится к $\hat{x}(t)$ при $\sigma \rightarrow 0$
2. $\frac{\partial x}{\partial \sigma}(t, \sigma)$ есть решение уравнения в вариациях

$$\dot{z} = \frac{\partial \varphi}{\partial x}(t, \hat{x}(t), \hat{u}(t))z$$

с начальными условиями $\frac{\partial x}{\partial \sigma}(\tau) = \varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau))$

Доказательство.

1. Следует из теоремы об интегральной непрерывной зависимости решений дифференциальных уравнений от параметра

2. Запишем уравнение (93) в интегральной форме

$$x(t, \sigma) = x_0 + \int_{t_0}^t \varphi(t, x(t, \sigma), \hat{u}(t))dt.$$

Поскольку $x(t, \sigma) \equiv \hat{x}(t)$ при $t < \sigma$, можно записать

$$x(t, \sigma) = \hat{x}(\tau) + \int_{\tau}^{\tau+\sigma} \varphi(t, x(t, \sigma), v)dt + \int_{\tau+\sigma}^t \varphi(t, x(t, \sigma), \hat{u}(t))dt.$$

Используя формулу дифференцирования интеграла с подвижными концами, получаем

$$\frac{\partial x}{\partial \sigma}(t, \sigma) = \varphi(\tau, \hat{x}(\tau), v) - \varphi(\tau, \hat{x}(\tau), \hat{u}(\tau)) + \int_{\tau}^{t_1} \frac{\partial \varphi}{\partial x}(t, \hat{x}(t), \hat{u}(t)) \frac{\partial x}{\partial \sigma}(t, \sigma) dt.$$

□

Лемма об интегрировании функционала

Лемма 10.

$$\frac{d}{d\sigma} \int_{t_0}^{t_1} f(t, x(t, \sigma), u(t, \sigma))dt = f(\tau, \hat{x}(\tau), v) - f(\tau, \hat{x}(\tau), \hat{u}(\tau)) + \int_{\tau}^{t_1} \frac{\partial x}{\partial \sigma}(t, \sigma) f_x(t, \hat{x}(t), \hat{u}(t)) dt.$$

Доказательство следует из теоремы о дифференцировании интеграла с переменными пределами с использованием леммы об иголке.

□

ДОКАЗАТЕЛЬСТВО принципа максимума Понтрягина.

Пусть $\hat{x}(t)$, $\hat{u}(t)$ оптимальный процесс, τ точка непрерывности управления $u(t)$. Возьмем вариацию типа иголки в точке τ . По лемме о дифференцировании функционала имеем

$$\frac{dL}{d\tau} = f(\tau, \hat{x}(\tau), v) - \hat{f}(\tau) + \int_{\tau}^{t_1} \frac{\partial x}{\partial \sigma}(t, \sigma) f_x(t, \hat{x}(t), \hat{u}(t)) dt. \quad (95)$$

Заметим, что \hat{f}_x есть свободный член в однородном линейном уравнении, которое является сопряженным к уравнению в вариациях для $z = \frac{\partial x}{\partial \sigma}$. Следовательно, интеграл, стоящий в правой части уравнения (95), можно заменить на левую часть уравнения (94). В результате, учитывая, что $p(t_1) = 0$, получим

$$\frac{dL}{d\tau} = f(\tau, \hat{x}(\tau), v) - \hat{f}(\tau) + p(\tau)(\phi(\tau, \hat{x}(\tau), v) - \hat{\phi}(\tau)).$$

Поскольку в точке минимума $\frac{dL}{d\tau} \geq 0$, все условия принципа максимума выполнены.

□

24 СУЩЕСТВОВАНИЕ РЕШЕНИЯ ЗАДАЧИ ОПТИМАЛЬНОГО БЫСТРОДЕЙСТВИЯ

В теории оптимального управления важную роль играют теоремы существования. В данной главе мы приведем примеры задач оптимального управления, в которых оптимального управления не существует. Значит ли это, что у рассматриваемых задач нет решения? Здесь уместно привести высказывание Гильберта: "Всякая задача вариационного исчисления имеет решение, если только слову "решение" придать соответствующий смысл". Это высказывание в равной мере применимо и к теории оптимального управления. Приведенные примеры указывают на пути обобщения понятия "решение" и позволяют пояснить смысл условий теоремы существования, которая доказывается во второй части данной главы.

Для того, чтобы ввести обобщенные решения, в вариационном исчислении и в оптимальном управлении пользуются понятием "минимизирующая последовательность".

Определение. Последовательность допустимых траекторий $\{x_n(\cdot)\}$ называется минимизирующей для задачи (88), если функционал $\mathcal{J}(x_n(\cdot))$ стремится к нижней грани $\mathcal{J}(x(\cdot))$ в классе всех допустимых траекторий.

Здесь под допустимой траекторией принимается траектория, соответствующая допустимому управлению.

Предел минимизирующей последовательности, если ему можно придать разумный смысл, принимается за определение обобщенного решения.

В дальнейшем нам понадобится следующее

Определение. Множество $f(x, U) = \bigcup_{u \in U} f(x, u)$ называется вектограммой.

§

12. Пример отсутствия оптимального управления. (Скольльзящие режимы)

Пример 3. Рассмотрим задачу наискорейшего попадания в начало координат из точки $(x(0) = -1, y(0) = 0)$ по траекториям системы

$$\begin{cases} \dot{x} = -y^2 + u^2 \\ \dot{y} = u \end{cases}, \quad |u| \leq 1 \quad (96)$$

Для того, чтобы доказать, что оптимального управления в этой задаче не существует, построим минимизирующую последовательность $\{x_n(\cdot), y_n(\cdot)\}$. Разделим отрезок, соединяющий точки $(-1, 0)$ и $(0, 0)$ на $2n$ равных частей. На всех нечетных участках возьмем $u = +1$, на всех четных $u = -1$. Получим последовательность пилообразных кривых, изображенных на рис. 11.

Для того, чтобы показать, что эта последовательность является минимизирующей, заметим, что проекция скорости движения на ось Ox не может быть больше 1: $\dot{x} = -y^2 + u^2 \leq 1$, причем эта

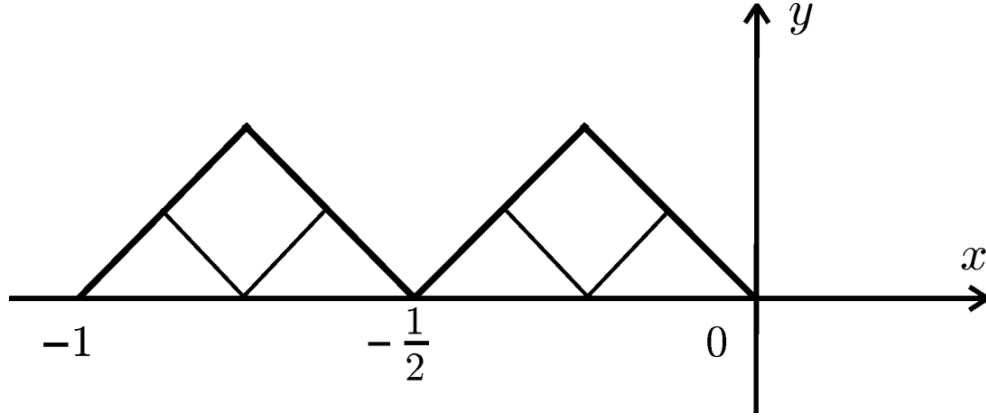


Рис. 11:

проекция будет тем ближе к единице, чем меньше $|y|$ и чем больше $|u|$. На последовательности $\{x_n(\cdot), y_n(\cdot)\}$ производная $\dot{x}_n(t)$ равномерно стремится к единице, так как $|u| = 1$ (п. в.), а $|y_n(t)|$ равномерно стремится к нулю. Отсюда следует, что время попадания в начало координат не менее единицы, и что на последовательности $\{x_n(\cdot), y_n(\cdot)\}$ оно стремится к единице. Покажем теперь, что за время, равное единице, попасть в начало координат нельзя. Предположим противное, тогда на соответствующей траектории $\dot{x} = 1$ (п. в.), и следовательно, $|u| = 1$ (п. в.), $y \equiv 0$, что противоречит уравнению $\dot{y} = u$. Итак, оптимального управления не существует.

Причина этого заключается в том, что множество $f(x, U)$ (которое в данном примере есть дуга параболы) невыпукло. Именно из-за этого траектория $y \equiv 0$, к которой сходится минимизирующая последовательность, не является решением системы (96) ни при каких допустимых управлениях, и, следовательно, множество допустимых траекторий некомпактно.

Разные компактификации множества допустимых траекторий приводят к различным понятиям обобщенного решения. Для задач с невыпуклой вектограммой обычно поступают следующим образом. Для управляемой системы

$$\dot{x} = f(x, u), \quad u \in U \quad (97)$$

рассматривается дифференциальное включение

$$\dot{x} \in \text{co } f(x, U), \quad (98)$$

где $\text{co } f(x, U)$ – выпуклое замыкание вектограммы $f(x, U)$. Решением дифференциального включения (98) называется абсолютно непрерывная функция $x(t)$, которая удовлетворяет включению (98) почти всюду.

Можно показать, что множество решений дифференциального включения (98) – компакт, и что множество допустимых траекторий системы (97) всюду плотно в этом компакте. Решения дифференциального включения (98), не являющиеся допустимыми траекториями системы (97), называются скользящими режимами для этой системы. Грубо говоря, скользящие режимы – это пределы таких траекторий, для которых управление на малых соседних участках принимает разные, чередующиеся друг с другом значения (см. рис. 11).

В рассматриваемом примере $\text{co } f(x, U)$ – параболический сегмент, который содержит вектор скорости $\dot{x} = 1$, $\dot{y} = 0$, следовательно, траектория $y(\cdot) = 0$ – скользящий режим.

§

13. Продолжимость решений обыкновенных дифференциальных уравнений
Рассмотрим систему

$$\dot{x} = \varphi(t, x), \quad (99)$$

где $x \in \mathbb{R}^n$, $\varphi : [t_0, t_1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ непрерывно дифференцируема по x и измерима по t .

Теорема I (о продолжимости решений).

Пусть существует константа $C > 0$ такая, что при всех $x \in \mathbb{R}^n$, $t \in [t_0, t_1]$

$$\langle x = \varphi(t, x) \rangle \leq C(|x|^2 + 1). \quad (100)$$

Тогда решения системы (99) определены на всем отрезке $[t_0, t_1]$.

Доказательство. Введем переменную $y(t) = |x|^2 + 1$. Имеем $\frac{dy}{dt} = 2\langle x, \frac{dx}{dt} \rangle = 2\langle x, \varphi(t, x) \rangle \leq 2C(|x|^2 + 1) = 2Cy$, т. е. $\frac{dy}{dt} \leq 2Cy$. Если $y(0) = A$, то $y(t) \leq Ae^{2C(t-t_0)}$, т. е.

$$|x| \leq \sqrt{A}e^{C(t-t_0)}. \quad (101)$$

что и требовалось доказать.

Отметим, что если неравенство (100) выполнено при всех \mathbb{R}_+^1 , то из теоремы 2 следует неограниченная по времени продолжимость решений.

Покажем, что свойство продолжимости решений существенно при изучении вопросов существования оптимального управления.

§

14. Пример отсутствия оптимального управления. (Уход на бесконечность за конечное время)

Пример 4.

Рассмотрим задачу наибоыстрейшего попадания в начало координат по траекториям системы

$$\begin{cases} \dot{x} = y \\ \dot{y} = u \exp(y^2) \end{cases}, \quad |u| \leq 1. \quad (102)$$

Применим принцип максимума $H = \psi_1 y + \psi_2 u \exp(y^2)$.

$$\dot{\psi}_1 = 0, \quad \dot{\psi}_2 = -\psi_1 - 2y\psi_2 \exp(y^2)u. \quad (103)$$

Из условия максимума имеем $u = 1$ при $\psi_2 > 0$; $u = -1$ при $\psi_2 < 0$ и любое при $\psi_2 = 0$.

Покажем, что $u(t)$ – управление, удовлетворяющее принципу максимума Понтрягина, имеет не более одного переключения. Действительно, переключение возможно только там, где $\psi_2(t)$ обращается в ноль. В силу (103) $\psi_1 = \text{const}$.

Пусть $\psi_1 = 0$, тогда уравнение для ψ_2 становится однородным. Покажем, что на соответствующей траектории не может быть ни одного переключения.

Действительно, если ψ_2 , обращается в ноль в некоторой точке, то по теореме единственности $\psi_2(t) \equiv 0$, следовательно, вектор ψ – нулевой, что противоречит принципу максимума.

Пусть $\psi_1 > 0$ тогда уравнение для ψ_2 имеет вид $\dot{\psi}_2 + a(t)\psi_2 = -\psi_1 < 0$. Покажем, что функция $\psi_2(t)$ не может сменить знак с минуса на плюс (и, тем самым, две перемены знака ψ_2 невозможны). Действительно, функция $a(t) = 2y(t)u(t) \exp(y^2(t))$ на любой допустимой траектории ограничена: $|a(t)| \leq C$. В области $|\psi_2| < \frac{\psi_1}{C}$ справедлива оценка $\dot{\psi}_2 = -\psi_1 - a(t)\psi_2 < 0$, т. е. в этой области ψ_2 убывает и поэтому не может сменить знак с минуса на плюс.

Аналогично показывается, что если $\psi_1 < 0$, то $\psi_2(t)$ не может сменить знак с плюса на минус.

Процесс построения синтеза проводится так же как в примере 1. Берем $u = 1$. Тогда система (102) принимает вид: $\dot{x} = y$, $\dot{y} = \exp(y^2)$. Разделив второе уравнение на первое и взяв x в

качестве независимого переменного, получим уравнение $\frac{dy}{dx} = \frac{\exp(y^2)}{y}$, траектории которого имеют вид $-\frac{1}{2} \exp(-y^2) = x + C$. Движение по этим траекториям происходит вверх, т. к. $\dot{y} > 0$.

Аналогичным образом, взяв $u = -1$, получим траектории $\frac{1}{2} \exp(-y^2) = x + C$. Движение по этим траекториям происходит вниз, т. к. $\dot{y} < 0$.

Скомбинировав из этих двух семейств траектории, ведущие в начало координат и имеющие не более одного переключения, получим:

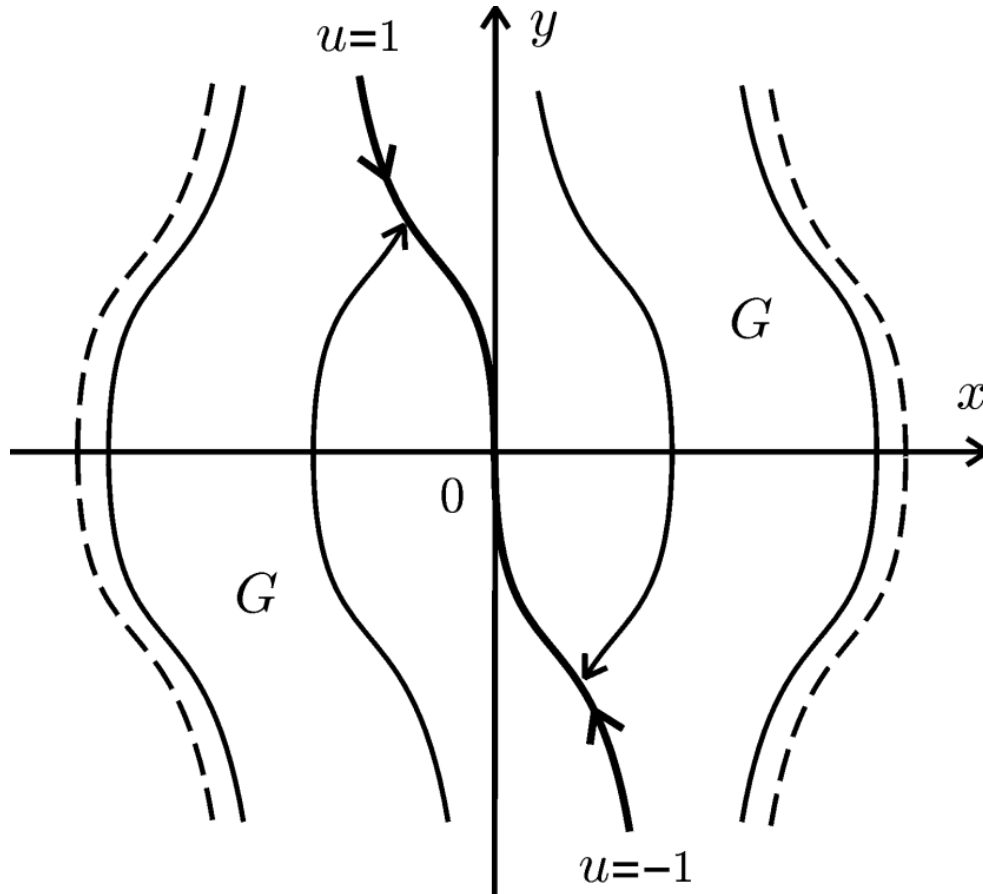


Рис. 12:

Здесь кривая переключения (изображенная на рис. 12 жирной линией) состоит из двух сопряженных друг с другом кривых: верхней половины кривой $x = \frac{1}{2} \exp(-y^2) - \frac{1}{2}$ и нижней половины кривой $x = -\frac{1}{2} \exp(-y^2) + \frac{1}{2}$. Область G , заполненная траекториями, лежит между кривыми $x = \frac{1}{2} \exp(-y^2) + \frac{1}{2}$ и $x = -\frac{1}{2} \exp(-y^2) - \frac{1}{2}$, изображенными на рис. 12 пунктиром. Для любой точки области G существует и притом только одна траектория, удовлетворяющая принципу максимума и ведущая в начало координат. Для точек, лежащих вне этой области, не существует ни одной траектории, удовлетворяющей принципу максимума и ведущей в начало координат.

Можно подумать, что для точек, лежащих вне области G не существует ни одной допустимой траектории системы (102), ведущей в начало координат. Однако это не так. Например, если для точки (x_0, y_0) , лежащей во втором ортанте, взять $u = 0$, то эта точка начнет двигаться вправо со скоростью y_0 и через конечное время достигнет области G . Следовательно, из точек, лежащих вне области G , нельзя попасть в начало координат оптимальным образом. Для того, чтобы осознать этот феномен полезно посмотреть на последовательность траекторий, минимизирующую

время перехода. Поскольку время перехода из точки (x_0, y_0) в начало координат на всех траекториях положительно, существует точная нижняя грань таких времен $T(x_0, y_0)$, и можно выбрать последовательность траекторий $\{x_n(\cdot), y_n(\cdot)\}$ начинающихся в данной точке (x_0, y_0) , такую, что время движения по ним до начала координат стремится к $T(x_0, y_0)$. Для точек, лежащих левее области G , траекторию $x_n(\cdot), y_n(\cdot)$ можно описать так: Сначала $u = 1$, при этом движение происходит вверх до уровня $y = n$, далее $u = 0$, при этом движение происходит по горизонтали до кривой переключения, затем $u = 1$, при этом движение происходит по кривой переключения до начала координат.

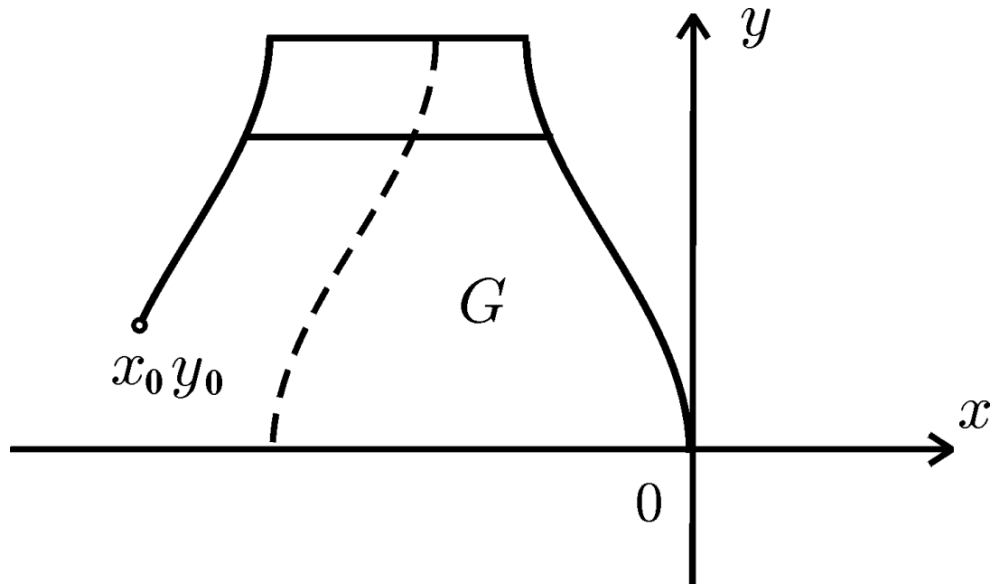


Рис. 13:

Справа от области G картина центрально симметрична.

Упражнение. Доказать, что вышеописанная последовательность траекторий действительно является минимизирующей.

Указание: а) времена движения по горизонтальным участкам траекторий $\{x_n(\cdot), y_n(\cdot)\}$, ($u = 0$) стремятся к нулю, т. к. длина всех этих участков ограничена величиной, не зависящей от n , а скорость движения по ним $\dot{x} = y = n \rightarrow \infty$; б) при $u = 1$ траектории системы (100) уходят на бесконечность за конечное время (именно этот факт является причиной несуществования оптимального управления в примере 4).

Замечание 1. Минимизирующая последовательность не стремится ни к какой допустимой траектории в обычном смысле этого слова. Однако, в данном случае можно ввести обобщенную предельную траекторию. Для точек, лежащих левее G , эта траектория поднимается ($u = +1$) до $+\infty$ и "оттуда" спускается по кривой переключения ($u = -1$) в начало координат. Формально говоря, эта траектория лежит в "расширенной" фазовой плоскости, компактифицированной двумя точками $y = +\infty$ и $y = -\infty$. В такой компактифицированной плоскости оптимальное управление существует.

Тема для самостоятельного исследования.

Изучить виды компактификации фазового пространства, связанные с различными способами ухода на бесконечность по траекториям системы (97). Сформулировать и доказать теорему существования оптимальных траекторий в компактифицированном фазовом пространстве.

§

15. Формулировка теоремы существования
Рассмотрим задачу быстродействия $T \rightarrow \inf$.

$$\dot{x} = f(x, u), \quad x(0) = a, \quad x(T) = b. \quad (104)$$

где $x \in \mathbb{R}^n$, $u \in U \subset \mathbb{R}^r$, $f: \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$.

В качестве класса допустимых управлений возьмем класс измеримых функций со значениями из множества U . Под решением системы (96) будем понимать решение по Каратеодори (см. [16], с.54).

Введем обозначение $f(x, U) = \bigcup_{u \in U} f(x, u)$.

Предположим, что выполнены следующие условия.

1. Функции $f(x, u)$, непрерывны по совокупности аргументов и непрерывно дифференцируемы по x .
2. Существует такая положительная константа C , что при всех $x \in \mathbb{R}^n$, $u \in U$

$$\langle x, f(x, u) \rangle \leq C(|x|^2 + 1).$$

3. Множество U замкнуто и ограничено; $f(x, U)$ выпукло.

Теорема 2 (теорема существования), (Филиппов). Пусть выполнены условия 1, 2, 3 и пусть существует допустимое управление $\tilde{u}(t)$ такое, что соответствующая ему траектория $\tilde{x}(t)$ с начальным условием $\tilde{x}(0) = a$ попадает в точку b при некотором $\tilde{t} > 0$: $\tilde{x}(\tilde{t}) = b$. Тогда существует оптимальное управление $u(t)$.

Прежде чем приступить к доказательству теоремы обсудим ее формулировку. Условие 1 играет чисто техническую роль. Его можно ослабить и оно действительно ослаблялось в целом ряде работ.

Условие 2 – это по существу условие теоремы 1. Для любого фиксированного отрезка времени $[0, T]$ оно гарантирует равномерную ограниченность множества решений системы (104), начинающихся в точке a и соответствующих любым измеримым управлениям $u(t)$. Как видно из примера 4, отказ от условия 2 может привести к несуществованию оптимального управления.

Условие 3 – это условие замкнутости множества решений системы (104), или, что то же, условие отсутствия скользящих режимов. Роль этого условия в вопросе существования оптимального управления выяснена в примере 3. Одновременное выполнение условий 2 и 3 гарантирует компактность множества траекторий системы (104), начинающихся в точке a .

§

16. Доказательство теоремы существования

Все рассуждения будут проводиться для отрезка $[0, \tilde{t}]$, на котором определена траектория $\tilde{x}(t)$. В силу оценки (101) при любом допустимом управлении $u(\cdot)$, $|x(t)|$ ограничен одной и той же константой: $|x| \leq \sqrt{A}e^{Ct}$. Поскольку функция $f(x, u)$ непрерывна на компакте $u \in U$, $|x| \leq \sqrt{A}e^{C\tilde{t}}$, имеем $|f(x, u)| \leq M$. Рассмотрим множество всех решений системы (104) (при допустимых $u(t)$, для которых $x(0) = a$, $x(t') = b$, $0 < t' \leq \tilde{t}$). Это множество не пусто, т. к. оно содержит $\tilde{x}(\cdot)$. Если это множество конечно, то утверждение теоремы очевидно. Если оно бесконечно, то выберем из него последовательность $x_n(\cdot)$ такую, что соответствующие времена перехода t'_n стремятся к $t_1 = \inf t'$. Функции $x_n(\cdot)$ равномерно ограничены, т. к. $|x_n(t)| \leq \sqrt{A}e^{Ct}$, и равномерно непрерывны, и почти всюду выполнено неравенство

$$|\dot{x}_n(t)| = |f(x_n(t), u_n(t))| \leq M \quad (\text{п.в.}) \quad (105)$$

По теореме Арцела из этой последовательности можно выбрать равномерно сходящуюся подпоследовательность, которую мы попрежнему будем обозначать $x_n(\cdot)$. Обозначим ее предел через

$x(\cdot)$. Покажем, что $x(\cdot)$ – решение нашей задачи. Поскольку t_1 есть нижняя грань времени перехода, не существует траектории, которая бы переводила a в b за время, меньшее t_1 . Для траектории $x(\cdot)$ имеем $x(0) = a$, $x_n(t'_n) = b$, $t'_n \rightarrow t_1$. В силу (105) $|x_n(t_1) - x_n(t'_n)| \leq M|t_1 - t'_n|$, поэтому $x(t_1) = b$.

Остается показать, что $x(t)$ – допустимая траектория; т. е. что $x(t)$ соответствует некоторому допустимому управлению. Именно эта часть доказательства является наиболее тонкой. Нам надо показать, что существует измеримое управление $u(t)$ такое, что $\dot{x}(t) = f(x(t), u(t))$. Заметим, во-первых, что поскольку все $x_n(t)$ удовлетворяют условию Липшица с одной и той же константой M , то и предельная траектория $x(t)$ обладает тем же свойством. Следовательно, $x(t)$ абсолютно непрерывна и $|\dot{x}| \leq M$ (п.в.).

Обозначим $\dot{x}(t) = y(t)$; $\dot{x}_n(t) = y_n(t)$. Функции $y_n(t)$ и $y(t)$ определены почти всюду на $[0, t_1]$, измеримы и ограничены. Пусть $t_0 \in [0, t_1]$ такая точка, в которой $\dot{x}(t_0)$ существует. Покажем, что $\dot{x}(t_0) \in f(x(t_0), U)$. В силу непрерывности функции f многозначное отображение $x \mapsto f(x, U)$ полунепрерывно сверху относительно включения, т. е. для любого $\epsilon > 0$ найдется такое $\delta > 0$, что как только $|x - x(t_0)| < \delta$, так $f(x, U) \in V_\epsilon$, где V_ϵ – замкнутая ϵ -окрестность множества $f(x(t_0), U)$. Уменьшив при необходимости δ , из дифференцируемости $x(\cdot)$ в точке t_0 получим

$$\left| \frac{x(t) - x(t_0)}{t - t_0} - \dot{x}(t_0) \right| < \epsilon \quad \text{при} \quad |t - t_0|, \delta. \quad (106)$$

Но $\frac{x(t) - x(t_0)}{t - t_0} = \lim_{n \rightarrow \infty} \frac{x_n(t) - x_n(t_0)}{t - t_0} = \lim_{n \rightarrow \infty} \frac{1}{t - t_0} \int_{t_0}^t y_n(\tau) d\tau = \lim_{n \rightarrow \infty} \int_0^1 y_n(t_0 + (t - t_0)s) ds$. Траектории $x_n(t)$ допустимы и $y_n(\tau) = \dot{x}_n(\tau) = f(x_n(\tau), u_n(\tau))$. Эти значения будут принадлежать V_ϵ , если $x_n(\tau)$ окажется достаточно близко к $x(t_0)$. Формальное доказательство этой близости проводится следующим образом: при $|\tau - t_0| < \delta$ и для всех достаточно больших n имеем $|x_n(t_0) - x(t_0)| < M\delta$, $|x_n(\tau) - x(\tau)| < M\delta$, откуда $|x_n(\tau) - x(t_0)| < 2M\delta$. Итак, $y_n(\tau) \in V_\epsilon$. Но V_ϵ – выпуклое замкнутое множество, как следует из условия 3. Поэтому интеграл $\int_0^1 y_n(t_0 + (t - t_0)s) ds$ также принадлежит V_ϵ , как предел выпуклых линейных комбинаций значений $y_n(\tau)$. (Обратите внимание, как красиво используется условие выпуклости!). Следовательно, предел этого интеграла, т. е. $\frac{x(t) - x(t_0)}{t - t_0}$ также принадлежит V_ϵ . Из (106) следует, что $\dot{x}(t_0) \in V_{2\epsilon}$ при любом сколь угодно малом $\epsilon > 0$. Из того, что $f(x_0, U)$ замкнуто, следует, что $\dot{x}(t_0) \in f(x(t_0), U)$, т. е. существует $u(t_0)$ такое, что $\dot{x}(t_0) = f(x(t_0), u(t_0))$. Последний этап доказательства состоит в том, что функцию $u(t)$ можно выбрать измеримой (в общем случае она определена неоднозначно). Доказательство этого факта следует из леммы об измеримом выборе, которую мы приведем без доказательства.

Лемма (об измеримом выборе). Пусть функция $f(t, u)$ непрерывна, $f : \mathbb{R} \times U \rightarrow \mathbb{R}^n$. Пусть измеримая вектор-функция $y(t)$ такова, что $y(t) \in f(t, U)$. Тогда существует измеримая вектор-функция $u(t)$ такая, что $y(t) = f(t, u(t))$ (п.в.).

Доказательство см. в [3] с.430.

ЛИТЕРАТУРА

1. Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф. Математическая теория оптимальных процессов. М., Наука, 1983.
2. Хартман Ф. Обыкновенные дифференциальные уравнения. - М., Мир, 1970.
3. Янг Л. Лекции по вариационному исчислению и теории оптимального управления. - М., Мир, 1974.
4. Никольский С. М. Курс математического анализа, тт. 1,2 - М., Наука, 1975.
5. Дубровин Б. А. Новиков С. П., Фоменко А. Т. Современная геометрия. - М., Наука, 1979.

6. Рашевский П. К. Курс дифференциальной геометрии. - М., ГИТТЛ, 1956.
7. Алексеев В. М., Тихомиров В. М., Фомин С. В. Оптимальное управление. - М., Наука, 1979.
8. Арнольд В. И. Математические методы классической механики. - М., Наука, 1974.
9. Блисс Г. А. Лекции по вариационному исчислению. - М., ИЛ, 1950.
10. Фоменко А. Т. Вариационные методы в топологии. - М., Наука, 1982.
11. Милнор Дж. Теория Морса. - М., Мир, 1965.
12. Алексеев В. М., Галеев Э. М., Тихомиров В. М. Сборник задач по оптимизации. - М., Наука, 1984.
13. Дьедонне Ж. Основы современного анализа. - М., Мир, 1964.
14. Ленг С. Алгебра. - М., Мир, 1968.
15. Нарасимхан Р. Анализ на действительных и комплексных многообразиях. М., Мир, 1971.
16. Коддингтон Э. Д., Левинсон Н. Теория обыкновенных дифференциальных уравнений. - М., ИЛ, 1958.
17. Филиппов А. Ф. Дифференциальные уравнения с разрывной правой частью. - М., Наука, 1985.
18. Лаврентьев М. А., Шабат Б. В. Методы теории функций комплексного переменного. М., Наука, 1973.