Вариационное исчисление и оптимальное управление (лекция 09.09.2014)

А.С. Демидов

МГУ, мех-мат.

1. Задача Иоганна Бернулли о брахистохроне и геодезические на римановой поверхности. Пусть требуется найти геодезические, т.е. (geo=землю daisia=разделяющие) кратчайшие линии на римановой поверхности, идущие от одного до другого объекта (например, от одной заданной точки до другой точки или до кривой). Если риманова метрика в декартовых координатах (t, x) такова, что

$$ds^2 = \frac{dt^2 + dx^2}{v^2(x)}\,, \qquad \text{где} \quad v(x) = (x + x_0)^\alpha \geq 0\,, \quad x_0 = \text{const} \in \mathbb{R}\,, \quad \alpha = \text{const} \in \mathbb{R}\,,$$

а объектами, соединяемыми искомыми геодезическими, являются либо две точки $(0, a_0)$ и $(0, a_1)$, либо точка $(0, a_0)$ и прямая $t = 1$, то речь идет о том, чтобы среди функций класса $C^1(0,1) \bigcap C[0,1]$, подчиненных условию

$$x(0) = a_0\,, \quad F(x) \stackrel{def}{=} \int_0^1 f(t,x,y)\Big|_{x=x(t),\, y=\dot{x}(t)} dt < \infty\,, \quad \text{где} \quad f(t,x,y) = \frac{\sqrt{1+y^2}}{(x+x_0)^\alpha}\,, \tag{1}$$

а в случае закрепленного правого конца еще и условию $x(1) = a_1$, найти такую функцию $\widehat{x}$, для которой

$$F(\widehat{x}) \leq F(x)\,, \qquad \text{что коротко записывают так:} \qquad F(x) \to \inf\,. \tag{2}$$

Если $\alpha = 0$ (евклидова геометрия), то ответ всем известен: геодезические — это прямые. Изящную интерпретацию получил случай $\alpha = 1/2$, $x_0 = 0$, исследование которого дало мощный импульс развитию математики, механики, физики. Все началось с того, что в 1696 году 29-летний Иоганн Бернулли (1667-1748), этот выдающийся представитель математической династии Бернулли, сформулировал, решил и, не раскрывая найденного им решения, предложил другим математикам решить *задачу о брахистохроне*[1], т.е. о форме такого жёлоба, задаваемому графиком функции

$$\widehat{x} : [0,1] \ni t \mapsto \widehat{x}(t)\,,$$

по которому под действием силы тяжести тело из точки $(0,0)$ за минимальное (гр. *brachistos* кратчайший) время (*chronos*) достигнет точку $(1, a_1)$, либо "стенку", задаваемую уравнением $t = 1$ (это так называемая задача со свободным правым концом). Еще в начале XVII века Галилей установил, что скорость v свободного падения пропорциональна корню квадратному пройденной высоты, т.е. $v(x) = C\sqrt{x}$. Ну а $\sqrt{1 + \dot{x}^2(t)}\, dt$ — это (согласно теореме Пифагора) длина элемента дуги жёлоба.

Будем исходить из того, что такая искомая функция $\widehat{x}$ существует[2]. Обозначим через $C_0^\infty(0,1]$ пространство тех функций h из $C^\infty(0,1]$, которые равны нулю, если $t \in [0, \varepsilon_h]$, где число $\varepsilon_h > 0$ зависит от h. Заметим, что $F(x) < \infty$, если $x = \widehat{x} + \lambda h$, где $h \in C_0^\infty(0,1]$ — любая фиксированная функция, а λ — достаточно малое по модулю число. А если $h \in C_0^\infty(0,1)$, т.е. функция h бесконечно гладкая и имеет компактный носитель, то будет выполнено и условию $x(1) = a_1$. Поэтому для $x = \widehat{x} + \lambda h$ имеем:

$$0 \stackrel{(2)}{\leq} F(x) - F(\widehat{x}) = \int_{\varepsilon_h}^1 [f(t, a+\alpha, b+\beta) - f(t,a,b)]\, dt\,, \qquad \text{где} \quad a = \widehat{x}(t),\ \alpha = \lambda h(t),\ b = \dot{\widehat{x}}(t),\ \beta = \dot{h}(t)\,. \tag{3}$$

[1]Задача о брахистохроне была инициирована таким вопросом Галилея: "по дуге $\smile AB$ окружности или по стягивающей ее хорде $[AB]$ тело быстрее спустится под действием силы тяжести из одной крайней точки в другую?" Галилей нашел ответ: быстрее по более длинному пути, т.е. по дуге окружности. А быстрее всего при какой форме жёлоба? Это уже задача о брахистохроне. Ее решение вскоре вслед за И.Бернулли нашли (опираясь на новые плодотворные идеи) Лейбниц, Ньютон, Лопиталь и его старший брат Якоб (с именем которого связано дифференциальное уравнение, лемниската и многочлен Бернулли, числа, распределение Бернулли и первая формулировка закона больших чисел).

[2]Предположение о существовании решения задачи на экстремум надо, вообще говоря, обосновывать. Иначе можно получить ошибочный результат. Например, предположение о существовании максимального натурального числа N сразу приводит к тому, что такое число $N = 1$. Действительно, если $N > 1$, то $N^2 - N = N(N-1) > 0$, т.е. $N^2 > N$, что невозможно, ибо N (как было предположено) есть самое большое натуральное число.

При любом фиксированном $t > \varepsilon_h$ числа a и $a+\alpha$ больше некоторого $\delta > 0$. А поскольку при $x > 0$ функция $f : (t,x,y) \mapsto f(t,x,y) = \frac{\sqrt{1+y^2}}{\sqrt{x}}$ принадлежит классу C^2 (и даже бесконечно дифференцируема), то

$$f(t, a+\alpha, b+\beta) - f(t,a,b) = [f'_x(t,a,b)\alpha + f'_y(t,a,b)\beta] + \frac{1}{2!}\Big[f''_{xx}(t,\tilde{a},\tilde{b})\alpha^2 + 2f''_{xy}(t,\tilde{a},\tilde{b})\alpha\beta + f''_{yy}(t,\tilde{a},\tilde{b})\beta^2\Big],$$

где $\tilde{a} = a + \theta\alpha$, $\tilde{b} = b + \theta\beta$, а $\theta \in (0,1)$. Таким образом, вводя обозначение

$$\widehat{g}(t) \stackrel{def}{=} g(t, \widehat{x}(t), \dot{\widehat{x}}(t)),$$

мы можем, учитывая то, что $h(t) = \dot{h}(t) = 0$ при $t \in [0, \varepsilon_h]$, переписать неравенство (3) в виде

$$0 \le A \cdot \lambda + B(\lambda), \tag{4}$$

где

$$A = \int_{\varepsilon_h}^1 [\widehat{f}'_x(t)h(t) + \widehat{f}'_{\dot{x}}(t)\dot{h}(t)]\,dt = \int_0^1 [\widehat{f}'_x(t)h(t) + \widehat{f}'_{\dot{x}}(t)\dot{h}(t)]\,dt$$

а $|B(\lambda)| \le \mathrm{const}\,\lambda^2$. Последняя оценка верна, т.к. при любом $\theta \in [0,1]$ функции $t \mapsto f''_{xx}(t,\tilde{a},\tilde{b})$, $t \mapsto f''_{xy}(t,\tilde{a},\tilde{b})$, $t \mapsto f''_{yy}(t,\tilde{a},\tilde{b})$ (а потому и интегралы от них) ограничены. Неравенство (4), которое верно при любом знаке малого параметра λ, влечет равенство $A = 0$. Итак,

$$\int_0^1 [\widehat{f}'_x(t)h(t) + \widehat{f}'_{\dot{x}}(t)\dot{h}(t)]\,dt = 0 \tag{5}$$

для любой функции $h \in C_0^\infty(0,1)$ в случае задачи с фиксированными концами, т.е. такой задачи

$$F(x) \to \inf, \qquad x \in \mathcal{M}_1 = \{\, x \in C^1(0,1), \quad x(0) = 0, \quad x(t) > 0 \quad \text{при} \quad t > 0 \quad \text{и} \quad x(1) = a_1 \,\} \tag{6}$$

и для любой функции $h \in C_0^\infty(0,1]$ в случае задачи со свободным правым концом, т.е. задачи

$$F(x) \to \inf, \qquad x \in \mathcal{M}_2 = \{\, x \in C^1(0,1), \quad x(0) = 0, \quad x(t) > 0 \quad \text{при} \quad t > 0 \,\}. \tag{7}$$

Условие (5) представляет собой бесконечную, несчетную (параметризованную функциями h) систему интегральных соотношений относительно искомой функции $\widehat{x}$. Но как “выудить” из бесконечного числа соотношений (5) ту информацию, которая относится непосредственно лишь к искомой функции $\widehat{x}$? Подсказкой может служить то, что известно с детства: температура T в точке ξ_0 характеризуется бесконечной последовательностью измерений $h_m \mapsto \int T(\xi)\, h_m(\xi - \xi_0)\, d\xi$ температуры в объемах, стягивающихся к точке ξ_0 с помощью все более и более локализованных около точки ξ_0 термометров (задаваемых сходящейся к δ-функции последовательностью $\{h_m\}_{m\ge1}$). В частности, предполагая функцию T непрерывной, получаем такую импликацию

$$\int_0^1 T(t)h(t)\,dt = 0 \quad \forall\, h \in C_0^\infty(0,1) \qquad \Rightarrow \qquad T(t) = 0 \quad \forall\, t \in (0,1).$$

Это совсем очевидное утверждение[3] носит название основная лемма вариационного исчисления.

Вернувшись к равенствам (5), мы должны преодолеть одну небольшую трудность. Дело в том, что интегральное слагаемое в (5) содержит и h, и $\dot{h}$. Можно последовать Лагранжу и, взяв по частям интеграл $\int_0^1 \widehat{f}'_{\dot{x}}(t)\dot{h}(t)\,dt$, избавиться от $\dot{h}$. Но для этого надо быть уверенным, что дифференцируема функция $\widehat{f}'_{\dot{x}} : t \mapsto \widehat{f}_{\dot{x}}(t)$. Априори уверенности в этом нет. В самом деле, если $f(t,x,\dot{x}) = \dot{x}^2$, а $x \in C^1$, то функция $t \mapsto f_{\dot{x}}(t)$ не дифференцируема. Тем не менее, как впервые (в 1879г.) показал Поль Дю Буа-Реймон (Du Bois-Reymond)[4] (1831-1889), справедлива

[3] Если в какой-то точке $t_0 \in (0,1)$ функция T отлична от нуля, то, взяв функцию $h \in C_0^\infty(0,1)$, равную нулю вне малой окрестности t_0, в которой h и непрерывная функция T сохраняет знак $T(t_0)$, получил отличное от нуля число $\int_0^1 T(t)h(t)\,dt$. А это противоречит условию $\int_0^1 T(t)h(t)\,dt = 0$.

[4] Поль Дю Буа-Реймон также впервые (в 1876) построил пример непрерывной функции, ряд Фурье которой расходится в некоторой точке. Будучи в окружении Вейрштрасса, он работал над проблемами теории функций вещественного переменного. На его работы ссылался Г. Кантор (в частности, потому, что Поль Дю Буа-Реймон вплотную подошел к методу диагонального процесса). Более известен старший брат Поля: основоположник нервно-мышечной физиологии Эмиль Дю Буа-Реймон (1818–1896). Их родители, имевшие французские корни (мать была дочерью французского представителя в Берлине) в 1804 году переехали в Берлин из швейцарского городка Нёшатель (фр. Neuchatel - “новый замок”).

Лемма 1 *Функция* $t \mapsto \widehat{f}_{\dot x}(t)$ *дифференцируема даже тогда, когда искомая функция*[5] $t \mapsto \widehat{x}(t)$ *априори всего лишь из класса* C^1.

♡ Чтобы это доказать, обозначим через R первообразную функции $\widehat{L}_x$. Тогда $\forall\, h \in C_0^\infty(0,1)$ имеем:

$$\int_0^1 \widehat{L}_x(t)h(t)\,dt = \int_0^1 \dot R(t)h(t)\,dt = -\int_0^1 R(t)\dot h(t)\,dt.$$

Следовательно, обозначая через G непрерывную функцию $G : t \mapsto G(t) \overset{def}{=} -R(t) + \widehat{L}_{\dot x}(t)$, получаем

$$0 \overset{(5)}{=} \int_0^1 \Big[\widehat{L}_x(t)h(t) + \widehat{L}_{\dot x}(t)\dot h(t)\Big]\,dt = \int_0^1 G(t)\dot h(t)\,dt \qquad \forall\, h \in C_0^\infty(0,1)\,.$$

В частности, для любых фиксированных точек t_1 и $t_2 \neq t_1$ из интервала $(0,1)$ имеем:

$$0 = \lim_{\varepsilon\to 0}\int_0^1 G(t)\dot h_\varepsilon(t)\,dt = \lim_{\varepsilon\to 0}\left[\int_0^1 G(t)\delta_\varepsilon(t-t_1)\,dt - \int_0^1 G(t)\delta_\varepsilon(t-t_2)\,dt\right] = G(t_1) - G(t_2)$$

где функция h_ε задается формулой (проверьте!, что $h_\varepsilon \in C_0^\infty(0,1)$)

$$h_\varepsilon(t) = \int_0^t \Big[\delta_\varepsilon(\tau - t_1) - \delta_\varepsilon(\tau - t_2)\Big]\,d\tau,$$

в которой $\delta_\varepsilon \geq 0$ — гладкая функция, равная нулю при $|t| > \varepsilon$ и такая, что $\int_{\mathbb{R}} \delta_\varepsilon(t)\,dt = 1$. Тем самым, функция G есть константа. Поэтому $\widehat{L}_{\dot x}(t) = R(t) + \mathrm{const} \in C^\infty$. Лемма доказана.

Итак, можно взять по частям интеграл

$$\int_0^1 \widehat{f}'_{\dot x}(t)\dot h(t)\,dt = -\int_0^1 \Big[\frac{d}{dt}\widehat{f}'_{\dot x}(t)\Big]h(t)\,dt + \widehat{f}'_{\dot x}(t)h(t)\Big|_0^1,$$

и переписать уравнения (5) в виде

$$\int_0^1 \Big[\widehat{f}'_x(t) - \frac{d}{dt}\widehat{f}'_{\dot x}(t)\Big]h(t)\,dt + \widehat{f}'_{\dot x}(1)h(1) = 0 \qquad \forall\, h \in C_0^\infty(0,1] \supset C_0^\infty(0,1) \tag{8}$$

в случае задачи (7) и в виде

$$\int_0^1 \Big[\widehat{f}'_x(t) - \frac{d}{dt}\widehat{f}'_{\dot x}(t)\Big]h(t)\,dt = 0 \qquad \forall\, h \in C_0^\infty(0,1) \tag{9}$$

в случае задачи (6).

Из последнего условия, которое справедливо для обеих задач, вытекает (согласно основной лемме вариационного исчисления), что искомая функция для каждой из этих задач должна быть решением дифференциального уравнения

$$\widehat{f}'_x(t) - \frac{d}{dt}\widehat{f}'_{\dot x}(t) = 0\,, \quad \text{т.е.} \quad f'_x(t,\widehat{x}(t),\dot{\widehat{x}}(t)) - \frac{d}{dt}\widehat{f}'_{\dot x}(t,\widehat{x}(t),\dot{\widehat{x}}(t)) = 0\,, \tag{10}$$

называемого уравнением Эйлера–Лагранжа[6] для функционала $F(x) \overset{def}{=} \int_0^1 f(t,x,y)\Big|_{x=x(t),\,y=\dot x(t)}\,dt.$

---

[5] В вариационных задачах $\int_{\Omega\ni(t,\xi)} L(t,\xi,x(t,\xi),x_t(t,\xi),x_\xi(t,\xi))\,dt\,d\xi \to \inf$ для функций многих переменных включение $\widehat{L}_{\dot x} \in C^1(\Omega)$ верно, вообще говоря, лишь если $\widehat{x} \in C^2(\Omega)$.

[6] Первоначально это уравнение получил Эйлер, аппроксимируя график искомой функции ломанной с шагом τ по переменой t и последующим предельным переходом при $\tau \to 0$ (т.е. рассматривая исходную экстремальную задачу для функционала как предельную задачу на минимум функции от конечного числа вещественных переменных — ординат вершин ломанной. В 1744 году вышел мемуар Эйлера о решении "изопериметрических задач в самом широком смысле", где он использовал тот же прием аппроксимации ломанными. А затем произошло следующее. В 1755 году Эйлер получил из Турина письмо от итальянца, ставшего впоследствии (как его прадед) французом Lagrange (1736–1813). Это было письмо 19-летнего Лагранжа, в котором он излагал свой взгляд на метод исследования изопериметрических задач. Какое удивительное совпадение: снова 19-летний юноша!, как и Эйлер, которому в 1726 году И. Бернулли предложил одну из

Общее решение этого уравнения 2-го порядка содержит два числовых параметра, которые в случае задачи (6) должны определяться из граничных условий $x(0) = a_0$, $x(1) = a_1$, а в случае задачи (7) из условия $x(0) = a_0$ и условия (8), принимающего (ввиду уравнения Эйлера–Лагранжа) следующий вид $\widehat{f}'_{\dot x}(1)h(1) = 0 \quad \forall\, h(1) \in \mathbb{R}$, т.е. из условия

$$\widehat{f}'_{\dot x}(1) = 0\,. \tag{11}$$

В отличие от граничного условия $x(0) = a_0$, граничное условие $\widehat{f}'_{\dot x}(1) = 0$ на правом конце интервала $(0, 1)$ не было заранее задано, а было найдено в результате интегрирования по частям интегральных тождеств (8) и равенства нулю возникших внеинтегральных слагаемых. Аналогичным образом находятся заранее не заданные граничные условия и в более общих задачах. В частности, в задаче на минимум $B(x) \to \inf$ для так называемого функционала Больца[7]

$$B : C^1[0,1] \ni x \mapsto \int_0^1 f(t, x(t), \dot x(t))\,dt + g(x(0), x(1)) \tag{12}$$

граничные условия определяются из условия

$$\widehat{f}'_{\dot x}(t)h(t)\Big|_0^1 + \frac{\partial g(\alpha,\beta)}{\partial \alpha}\Big|_{(\alpha,\beta)=(\widehat{x}(0),\widehat{x}(1))} h(0) + \frac{\partial g(\alpha,\beta)}{\partial \beta}\Big|_{(\alpha,\beta)=(\widehat{x}(0),\widehat{x}(1))} h(1) = 0\,, \tag{13}$$

где $h(0)$, соответственно, $h(1)$ равно нулю, если и только если задано граничное условие вида $x(0) = a_0$, соответственно, $x(1) = a_1$. Если же $h(0)$ и/или $h(1)$ может принимать любое значение (т.е. если явно не задано граничное условие при $t = 0$ и/или $t = 1$), то вытекающие из (13) граничные условия (какие?) показывают, как график искомой функции $\widehat{x}$ должен пересекать заданное концевое множество (такое, например, как прямая $\{t = 1\}$ в задаче (7)). Поэтому эти условия называют условиями трансверсальности.

В случае задачи о брахистохроне, $f(t, x, y) = \frac{\sqrt{1+y^2}}{\sqrt{x}}$. Заметим, что $\frac{\partial f(t,x,y)}{\partial t} \equiv 0$. Этот факт (как мы сейчас увидим) позволяет найти ту функцию $F : \mathbb{R}^2 \ni (x, y) \mapsto F(x, y)$, линии уровня которой, параметризованные константой C, т.е. линии $\{\ (x, y) \ni \mathbb{R}^2 \mid F(x, y) = C\ \}$ задают интегральные кривые уравнения Эйлера–Лагранжа. Такие функции называются первым интегралом дифференциального уравнения, поскольку соотношение $F(x, y) = C$ понижает на единицу порядок дифференциального уравнения, т.е. позволяет как бы один раз его проинтегрировать. В данном случае речь идет об автономной динамической системе, т.е. не зависящей от времени ($\frac{\partial f(t,x,y)}{\partial t} \equiv 0$). В таких системах энергия постоянна. Поэтому соответствующий первый интеграл, т.е. функция F называется интегралом энергии. В случае уравнения Эйлера–Лагранжа эта функция задается формулой

$$F(x, y) = \widehat{f}(t) - \dot{\widehat{x}}(t)\widehat{f}_{\dot x}(t) \qquad \text{т.е.} \qquad \widehat{f}(t) - \dot{\widehat{x}}(t)\widehat{f}_{\dot x}(t) = \mathrm{const}\,. \tag{14}$$

этих задач. Письмо содержало вывод уравнения (8), базировавшийся на сравнительном анализе исследуемого на минимум интеграла при возмущении (вариации) искомой функции $\widehat{x}$. (Кстати, по этой причине Эйлер ввел в науку термин “вариационное исчисление”.) Восхищенный идеей Лагранжа, Эйлер посылает ему восторженное письмо. Начинается оживленная переписка. Эйлер получает новые результаты. Но он не спешит их публиковать и 10 октября 1759 года посылает Лагранжу такое в высшей степени знаменательное письмо: “Твое аналитическое решение изопериметрической проблемы содержит, насколько я вижу, все, чего только можно желать в этой области, и я чрезвычайно рад, что эта теория, которой после моих первых попыток я занимался едва ли не один, доведена тобой до величайшего совершенства. Важность вопроса побудила меня к тому, что я с помощью твоего освещения сам вывел аналитическое решение. Я, однако, решил скрывать это, пока ты не опубликуешь свои результаты, так как я никоим образом не хочу отнимать у тебя часть заслуженной тобой славы.”

Часть этой славы в том, что уравнение (10) известно в науке, как *уравнение Эйлера–Лагранжа.*

[7]Оскар Больца (1857-1942), немецкий математик, сформулировал в 1913 г. задачу Лагранжа, где функционал имеет не только интегральную часть $\int f$ (что было у Лагранжа), но и терминальную (функцию от концов) g. Впрочем, вводя в задаче Лагранжа дополнительную искомую функцию y, подчиненную дополнительному дифференциальному уравнению $\dot y = 0$ и дополнительному терминальному ограничению $y(T) = g(\theta, T, x(\theta), x(T))$, получаем формулировку задачи Лагранжа с функционалами лишь интегрального вида. А вводя переменную $\varkappa$, подчиненную уравнению $\dot\varkappa_j = f(t, x, \dot x)$ и граничному условию $\varkappa(\theta) = 0$, интегральную часть функционала F можно представить в виде функции от концов, которую принято называть функционалом Майера в честь немецкого математика Адольфа Майера (1839-1908). Именно А. Майер в 1886 г. впервые сформулировал и показал (впрочем, далеко не в полной мере), что принцип Лагранжа для задачи Лагранжа верен, если искомых функций больше числа связывающих их дифференциальных уравнений 1-го порядка. Спустя 10 лет Д. Гильберт (см. его *Избранные труды*, том 2, “Факториал”, Москва 1998) дал строгое доказательство этого результата А. Майера, опираясь на непрерывную зависимость решений системы дифференциальных уравнений $\dot x = \varphi(t, x)$ от правой части. Отметим, что занимаясь вариационными задачами в ранге профессора Гейдельбергского университета (в котором, заметим, преподавали многие знаменитости, в том числе один из творцов немецкой классической философии Георг Гегель), Майер подошел вплотную к теории непрерывных групп Софуса Ли, с которым вел оживленную переписку.

Докажите утверждение (14), подправив следующую (стандартную, но не вполне корректную) выкладку

$$\frac{d}{dt}\Big(\widehat{f}(t)-\dot{\widehat{x}}(t)\widehat{f}_{\dot{x}}(t)\Big)=\widehat{f}_x(t)\dot{\widehat{x}}(t)+\widehat{f}_{\dot{x}}(t)\ddot{\widehat{x}}(t)-\ddot{\widehat{x}}(t)\widehat{f}_{\dot{x}}(t)-\dot{\widehat{x}}(t)\frac{d}{dt}\widehat{f}_{\dot{x}}(t)=\dot{\widehat{x}}(t)\Big(\widehat{f}_x(t)-\frac{d}{dt}\widehat{f}_{\dot{x}}(t)\Big)\overset{(10)}{\equiv}0\,,$$

показав при этом, что вторая производная функции $\widehat{x}$ существует там, где $\widehat{f}_{\dot{x}\dot{x}}(t)\neq 0$.

Указание. Хотя $\ddot{\widehat{x}}(t)$ может и не существовать, однако для $A(\tau)\overset{def}{=}\frac{f(\widehat{x}(t),\dot{\widehat{x}}(t+\tau))-f(\widehat{x}(t),\dot{\widehat{x}}(t))}{\tau}$ существует предел $\lim\limits_{\tau\to 0}A(\tau)=\frac{d}{dt}\widehat{f}(t)-\widehat{f}_x(t)\dot{\widehat{x}}(t)$. Отсюда и формулы Лагранжа для конечных разностей вытекает, что существует равный ему предел $\lim\limits_{\tau\to 0}B(\tau)$, где $B(\tau)\overset{def}{=}\frac{\left[\dot{\widehat{x}}(t+\tau)-\dot{\widehat{x}}(t)\right]f_{\dot{x}}(\widehat{x}(t+\tau),\dot{\widehat{x}}(t+\tau))}{\tau}$. А так как существует $\frac{d}{dt}\widehat{f}_{\dot{x}}(t)$, то существует $\frac{d}{dt}\dot{\widehat{x}}(t)\widehat{f}_{\dot{x}}(t)=\lim\limits_{\tau\to 0}B(\tau)+\dot{\widehat{x}}(t)\frac{d}{dt}\widehat{f}_{\dot{x}}(t)$. Далее, $\frac{d}{dt}\widehat{f}_{\dot{x}}(t)=f_{\dot{x}x}\big(\widetilde{x}(t),\dot{\widetilde{x}}(t)\big)\dot{\widehat{x}}(t)+\lim\limits_{\tau\to 0}f_{\dot{x}\dot{x}}\big(\widetilde{x}(t),\dot{\widetilde{x}}(t)\big)\frac{\dot{\widehat{x}}(t+\tau)-\dot{\widehat{x}}(t)}{\tau}$, где $\widetilde{x}(t)=\widehat{x}(t)+\theta\big(\widehat{x}(t+\tau)-\widehat{x}(t)\big)$, $\dot{\widetilde{x}}(t)=\dot{\widehat{x}}(t)+\theta\big(\dot{\widehat{x}}(t+\tau)-\dot{\widehat{x}}(t)\big)$, а $\theta\in(0,1)$. Отсюда $\exists\ddot{x}(t)$, если $\widehat{f}_{\dot{x}\dot{x}}(t)\neq 0$.

Найдем общее решение уравнения Эйлера–Лагранжа для задачи о брахистохроне, воспользовавшись наличием интеграла энергии (см. (14)), что в данном случае приводит к такому уравнению первого порядка

$$x=D/(1+\dot{x}^2)\,.$$

Учитывая, что $\dot{x}$ может априори принимать любые значения (и что, согласно предсказанию Галилея $\dot{x}(0)=\infty$) сделаем замену $\dot{x}=\mathrm{ctg}\varphi$ (тем более, что имеется красивая формула $1/(1+\mathrm{ctg}^2\varphi)=\sin^2\varphi$). Тогда

$$x=\frac{D}{1+\mathrm{ctg}^2\varphi}=D\sin^2\varphi=R(1-\cos 2\varphi)\,,\qquad \text{где}\quad R=D/2\,.$$

Кроме того,

$$\frac{dt}{d\varphi}=\frac{dx/d\varphi}{dx/dt}=\frac{2D\sin\varphi\cos\varphi}{\mathrm{ctg}\varphi}=2D\sin^2\varphi=R(2-2\cos 2\varphi)\qquad\Rightarrow\qquad t(\varphi)=R(2\varphi-sin2\varphi)+t_0\,.$$

Полагая $\theta=2\varphi$ и учитывая начальное условие $x(0)=0$, получаем такое параметрическое задание искомой функции

$$x=R(1-\cos\theta)\qquad t=R(\theta-sin\theta)\,.$$

График этой функции есть циклоида, т.е. кривая, которую описывает метка A на обруче радиуса R, при его повороте на угол θ в процессе движении вдоль оси абсцисс, если в начальный момент (при $\theta=0$) метка A находилась на оси абсцисс в точке $t=0$.

В задаче (6) с фиксированным правым концом, параметр R определяется из условия $\widehat{x}(1)=a_1$, а в случае задачи (7) со свободным правым концом, нужно брать циклоиду, которая ортогонально пересекает прямую $t=1$ (ибо условие трансверсальности (11) в данном случае означает, что $\dot{\widehat{x}}(1)=0$).

2.[8] Факторпространства, Замкнутые подпространства банахова пространства

В $\mathbb{R}^n$ любое его линейное подпространство замкнуто, т.е. содержит все свои предельные точки.

Однако в бесконечномерном гильбертовом или банаховом пространстве это не так. В частности, не замкнуто в X любое линейное подпространство $X_0\subsetneq X$, которое плотно в X, как например, подпространство $X_0=C^1$ в $X=C$ или в $X=L^2$. Примером замкнутого подпространств является, конечно, нуль-пространство $X_0=\{x\in X\mid Ax=0\}$ любого оператора $A\in L(X,Y)$, где

$$L(X,Y)\quad\text{есть } \textit{пространство линейных непрерывных операторов} \text{ из } X \text{ в } Y\,. \tag{15}$$

Нуль-пространство оператора $A\in L(X,Y)$ называют ядром этого оператора и обозначают так: $\mathrm{Ker}A$.

В дальнейшем нам будет важно знать, является ли замкнутым в Y образ оператора $A\in L(X,Y)$, т.е. множество $\mathrm{Im}A\overset{def}{=}\{\,y\in Y\mid y=Ax\,\}$. Обозначая через $\overline{\mathcal{M}}$ замыкание множества $\mathcal{M}$, т.е. множество всех его предельных точек, интересующий нас вопрос можно сформулировать так: верно ли, что $\mathrm{Im}A=\overline{\mathrm{Im}A}$? Это, конечно, не так, если $\mathrm{Im}A=Y_0$, где $Y_0\subsetneq Y$, а $\overline{Y}_0=Y$, скажем, если $Y_0=C^1$, а $Y=C$ или $Y=L^2$. Такая ситуация возникает, например, когда $X=Y_0$, а оператор A является идентичиым, т.е. оператором вложения Y_0 в Y. Другим примером является оператор "интегрирования" $A:C[0,1]\ni x\mapsto\int\limits_0^t x(\tau)\,d\tau\in C[0,1]$ или $A:l^2=h^0\ni(x_1,\dots,x_n,\dots)\mapsto(x_1,\dots,\frac{x_n}{n},\dots)\in l^2$. В первом из этих случаев $\mathrm{Im}A=C^1[0,1]\subsetneq C[0,1]=\overline{C^1[0,1]}$, а во втором $\mathrm{Im}A=h^1\subsetneq h^0=\overline{h^1}$[9]. В обоих случаях образ оператора $A\in L(Y,Y)$ является незамкнутым линейным подпространством в Y.

[8] На этот пункт, который предполагался в лекции 09.09, в тот день не хватило времени.

[9] h^m — пространство последовательностей $x=(x_1,\dots,x_n,\dots)$ с $\|x\|_m\overset{def}{=}\sum\limits_{n\geq 1}(n^m x_n)^2<\infty$.

В нижеследующей лемме 2, в которой приводится важное для нас достаточное условие замкнутости образа оператора $A \in L(X,Y)$, действующего в банаховых пространствах X и Y, фигурирует понятие фактор-пространства X/X_0 линейного пространства X по линейному подпространству $X_0 \subset X$. Это есть линейное пространство так называемых классов смежности $\widetilde{x}$, в котором нулевым элементом $\widetilde{0}$ считается пространство X_0, а любой другой элемент $\widetilde{x} \in X/X_0$ есть сдвиг $\widetilde{0} = X_0$ на некоторое $x \in X$ или любое другое $x' \in X$, но только такое, для которого $x' - x \in X_0$. Так например, два вектора (x_1, x_2, x_3) и $x_1', x_2', x_3')$ из координатного пространства $X = \mathbb{R}^3$ представляют один и тот же элемент в X/X_0, где $X_0 = \{\, (0,0,a) \mid a \in \mathbb{R}\,\}$, тогда и только тогда, когда $(x_1', x_2') = (x_1, x_2)$. Иными словами, $(x_1', x_2', x_3') - (x_1, x_2, x_3) \in X_0$. Вот другой пример: $C[0,1]/C^1[0,1]$ — это линейное пространство классов непрерывных функций $\widetilde{x}$, представимых в виде $x + \varphi$, где φ — любая функция из $C^1[0,1]$. Если $0 < a < b < 1$, а $x_a(t) = |t-a|$, $x_b(t) = |t-b|$, то предположив, что линейная комбинация элементов $\widetilde{x}_a$ и $\widetilde{x}_b$ равна нулю в $C[0,1]/C^1[0,1]$, иначе говоря, $\lambda_a \widetilde{x}_a + \lambda_b \widetilde{x}_b = \widetilde{0}$, получим: $\lambda_a = \lambda_b = 0$, т.е. элементы $\widetilde{x}_a$ и $\widetilde{x}_b$ линейно независимы в $C[0,1]/C^1[0,1]$. Отсюда следует, что $\dim\bigl(C[0,1]/C^1[0,1]\bigr) = \infty$. Этот частный результат обобщает весьма полезная

Лемма 2 *Пусть* $A \in L(X,Y)$ *и пусть* $\dim\,(Y/\mathrm{Im}A) < \infty$. *Тогда* $\mathrm{Im}A$ *– замкнут в* Y.

Доказательство. Так как $\dim\, Y/\mathrm{Im}A < \infty$, то $Y = \mathrm{Im}A \dot{+} M$ (прямая сумма линейных пространств), где $\dim M < \infty$, и потому M, снабженное некоторой нормой $\|\cdot\|_M$ есть банахово пространство. Рассмотрим оператор A_1, действующий из банахова[10] $X_1 = \bigl(X/\mathrm{Ker}A\bigr) \times M$ в банахово $Y = \mathrm{Im}A \dot{+} M$ по формуле: $A_1(\widetilde{x}, m) = Ax + m$. Обратный к A_1 существует, ибо A_1 имеет нулевое ядро. Ясно, что оператор A_1 непрерывен. Он действует "на". По теореме Банаха об обратном операторе[11], оператор $B = A_1^{-1}$ непрерывен. Отсюда вытекает, что $\mathrm{Im}\,A = B^{-1}(X/\mathrm{Ker}A \times \{0\})$ замкнут в Y, как прообраз замкнутого множества $X/\mathrm{Ker}A \times \{0\} \subset X_1$ при непрерывном отображении $B : Y = \mathrm{Im}A \dot{+} M \to X/\mathrm{Ker}A \times M$. $\square$

Лемма 2 имеет немало применений. В частности, она используется в нижеследующем доказательстве теоремы о принципе Лагранжа для гладких экстремальных задач.

Теорема 1 *Пусть* Z, Y *— банаховы пространства,* $F : Z \xrightarrow{C^1} \mathbb{R}$, $Q : Z \xrightarrow{C^1} Y$, *а* $\widehat{z}$ *— решение задачи*

$$F(z) \to \inf\,, \quad z \in \mathcal{M} = \{z \in Z \mid Q(z) = 0\ \}\,.$$

Пусть $\mathrm{Im}A = Y$, *где* $A = Q'(\widehat{z})$. *Тогда существует такой элемент* $y^* \in Y^* \stackrel{def}{=} L(Y,\mathbb{R})$, *что* $F'(\widehat{z}) + \langle y^*, Q'(\widehat{z})\rangle = 0$, *иначе говоря,*

$$\langle F'(\widehat{z}) + y^* Q'(\widehat{z}), \zeta\rangle = 0 \qquad \textit{для} \qquad \forall\, \zeta \in Z\,, \qquad \textit{где} \quad y^* Q'(\widehat{z}) \stackrel{def}{=} \langle y^*, Q'(\widehat{z})\rangle\,. \tag{16}$$

Доказательство. Введем оператор $D : Z \ni \zeta \mapsto D\zeta = \Bigl(F'(\widehat{z})\,\zeta\,,\, Q'(\widehat{z})\,\zeta\Bigr) \in Y_1 \stackrel{def}{=} \mathbb{R} \times Y$. Отождествляя те элементы $d_1 = (b_1, a_1)$ и $d_2 = (b_2, a_2)$ координатного пространства $\mathbb{R} \times Y$, для которых $(b_1, a_1) - (b_2, a_2) \in \mathrm{Im\,D}$, рассмотрим линейное пространство $Y_1/\mathrm{Im\,D}$. Имеем:

i) $\dim\bigl(Y_1/\mathrm{Im\,D}\bigr) \le 1$ (т.к. $\mathrm{Im\,Q}'(\widehat{z}) = Y$) и потому[12], **в силу леммы 2**, $\mathrm{Im\,D}$ замкнут в Y_1;

ii) $\mathrm{Im\,D} \neq Y_1$, т.к. $(1,0) \notin \mathrm{Im\,D}$, что верно, ибо иначе $\bigl(F'(\widehat{z})\,\zeta\,,\, Q'(\widehat{z})\,\zeta\bigr) = (1,0)$ и потому $\zeta \in \mathrm{Ker}A$, что влечет (в силу так называемой леммы Люстерника о касательном пространстве) противоречие: $F'(\widehat{z})\,\zeta = 0$.

Применяя теорему Хана–Банаха к замкнутому в $Y_1 = \mathbb{R} \times Y$ подпространству $\mathrm{Im\,D} \neq Y_1$, возьмем такой *ненулевой*(!) функционал $\Lambda = (\lambda, y^*) \in Y_1^*$, что $\langle \Lambda, y\rangle = 0 \ \ \forall\, y \in \mathrm{Im\,D}$. В частности, для $y := \bigl(F'(\widehat{z})\,\zeta\ Q'(\widehat{z})\,\zeta\bigr)$ $\langle \lambda F'(\widehat{z}) + y^* Q'(\widehat{z}), \zeta\rangle = 0$ для $\forall\, \zeta \in Z$. Осталось заметить, что $\lambda \neq 0$. Это так, ибо иначе $\langle y^* Q'(\widehat{z}), \zeta\rangle = 0$ для $\forall\, \zeta \in Z$, что дает $y^* = 0$, так как $\mathrm{Im}Q'(\widehat{z}) = Y$. Тем самым, предположение $\lambda = 0$ влечет $\Lambda = (\lambda, y^*) = 0$, вопреки тому, что мы взяли $\Lambda \neq 0$. $\square$

Примечание. Как показывает пример задачи $F(x,y) \stackrel{def}{=} y \to \inf$, $(x,y) \in \mathfrak{M}$, где $\mathfrak{M}$ есть замыкание множества $\{(x,y) \in \mathbb{R}^2 \mid Q(x,y) = 0\}$, а $Q(x,y) = x^2 \sin\frac{1}{x - \sqrt{y}}$, принцип Лагранжа справедлив без предположения о строгой дифференцируемости отображения Q.

---

[10]Норма в прямом произведении $\widetilde{X} \times M$, где $\widetilde{X} = X/\mathrm{Ker}A$ задается формулой $\|(\widetilde{x}, m)\|_{X_1} = \|\widetilde{x}\|_{\widetilde{X}} + \|m\|_M$. Здесь $\|\widetilde{x}\|_{\widetilde{X}} = \inf\limits_{a \in \mathrm{Ker}A} \|x + a\|_X$ есть норма (см., например, [КФ]) класса смежности $\widetilde{x} \in X/\mathrm{Ker}A$, представителем которого является элемент $x \in X$. Действительно, $\|\widetilde{0}\|_{\widetilde{X}} = 0$, $\|\widetilde{x} + \widetilde{y}\|_{\widetilde{X}} \le \|\widetilde{x}\|_{\widetilde{X}} + \|\widetilde{y}\|_{\widetilde{X}}$ и $\|a\widetilde{x}\|_{\widetilde{X}} = |a|\,\|\widetilde{x}\|_{\widetilde{X}}\ \forall\, a \in \mathbb{R}$. Эти свойства нормы очевидны. Наконец, если $0 = \|\widetilde{x}\|_{\widetilde{X}}$, то $0 = \inf\limits_{\mathrm{Ker}A \ni a_n \to a} \|x + a_n\|_X = \|x + a\|_X$. При этом $a = \lim\limits_{n\to\infty} \in \mathrm{Ker}A$ (ввиду замкнутости $\mathrm{Ker}A$), поэтому $x = -a \in \mathrm{Ker}A$, т.е. $\widetilde{x} = 0$. Снабженное нормой $\|\cdot\|_{\widetilde{X}}$, фактор-пространство $\widetilde{X}$ наследует полноту пространства X. В самом деле, пусть $\{\widetilde{x}_n\}$ — такая подпоследовательность фундаментальной последовательности в $\widetilde{X}$, что $\|\widetilde{x}_{n+1} - \widetilde{x}_n\|_{\widetilde{X}} \le 1/2^n$. Выберем представитель $x_1 \in X$ элемента $\widetilde{x}_1 \in \widetilde{X}$, а затем при $n \ge 1$ возьмем такой представитель $x_{n+1} \in X$ элемента $\widetilde{x}_{n+1} \in \widetilde{X}$, что $\|x_{n+1} - x_n\|_X \le 1/2^n$. Для любого $m \in \mathbb{N}$ имеем: $\|x_{n+m} - x_n\|_X \le 1/2^n(1 + 1/2 + \ldots + 1/2^m) \le 1/2^{n-1}$. В силу полноты X, существует такое $x \in X$, что $\|x_n - x\|_X \to 0$, а потому для $\widetilde{x} \in \widetilde{X}$ имеем: $\|\widetilde{x}_n - \widetilde{x}\|_{\widetilde{X}} \to 0$.

[11]Пусть X и Y – банаховы пространства, $A \in L(X,Y)$. Если $\mathrm{Ker}A = 0$, то $\exists A^{-1} : \mathrm{Im}A \to X$. Однако оператор A^{-1} может не быть непрерывным (например, если $A : (x_1, \ldots, x_n \ldots) \mapsto (x_1, \ldots, \frac{1}{n}x_n \ldots)$ — оператор "интегрирования" в l^2). Теорема Банаха об обратном операторе (см., например, [КФ]), утверждает, что A^{-1} будет непрерывным, если $\mathrm{Im}A = Y$.

[12]Пусть $C = (B, A) : l^2 \to l^2 \times l^2$, где $Bx = (x_1, 0, x_3, 0, x_5, 0, \ldots)$, а $Ax = (0, x_1 + \frac{x_2}{2}, 0, x_3 + \frac{x_4}{4}, 0, x_5 + \frac{x_6}{6}, 0, \ldots)$ Этот простой пример показывает, что образ оператора $C = (B, A) \in L(Z, X \times Y)$ может не быть замкнутым даже тогда, когда операторы $B \in L(Z,X)$ и $A \in L(Z,Y)$ имеют замкнутый образ. Однако, если $\mathrm{Im}B$ замкнут и на $\mathrm{Ker}A$ (как, например, $\mathrm{Im}\,F'(\widehat{z})$ на $\mathrm{Ker}\,Q'(\widehat{z})$), то $\mathrm{Im}C$ замкнут. Этот результат, вошедший во многие учебники по оптимальному управлению, принадлежит Владимиру Михайловичу Алексееву (1932 - 1980), профессору мех-мат факультета МГУ, крупнейшему специалисту по качественным методам небесной механики и символической динамики.