Лекция 25 (10).

Гл. X. Элементы вариационного исчисления.

§1. Основные понятия.

Пусть M – некоторое множество функций.

Определение. ***Функционалом*** $J = J[y]$ называется переменная величина, зависящая от функции $y(x)$, если каждой функции $y(x) \in M$ по некоторому правилу поставлено в соответствие число: $y(x) \to J[y(x)]$.

Определение. Множество M функций $y(x)$, на котором определен функционал $J[y]$, называется его ***областью определения***.

Таким образом, можно сказать, что функционал - это функция от функции: функция, в которой роль независимого переменного играет функция из некоторого множества M .

Примеры функционалов:

1) $J = f\left(x^0\right)$ – значение функции в точке $x^0 \in [a,\, b]$,
2) точная верхняя/нижняя грань функции $y(x) \in C([a,\, b])$,
3) $J = \int\limits_a^b y(x)dx$ – площадь под кривой,
4) $J = \int\limits_a^b \sqrt{1 + \left(y'(x)\right)^2}\, dx$ – длина кривой .

Наиболее разработанными в "исчислении функционалов" являются методы нахождения наибольших и наименьших значений функционалов, получившие название ***вариационного исчисления***.

Так же, как и для функций, для функционалов вводятся (почти дословно) такие понятия как

- непрерывность функционала,
- первый дифференциал функционала.

Функция $f(x)$ непрерывна в точке $x = x^0$, если $\forall \varepsilon > 0 \quad \exists \delta = \delta_\varepsilon > 0: \quad \forall x: \quad \left\lvert x - x^0 \right\rvert < \delta \quad \to \quad \left\lvert f(x) - f\left(x^0\right) \right\rvert < \varepsilon$, т.е. малому изменению аргумента соответствует малое изменение функции.	Функционал $J[y(x)]$ называется непрерывным, если малому изменению функции соответствует малое изменение функционала.

Остается уточнить, какие изменения функции считать малыми, т.е. какие кривые $y = y(x)$ и $y = y_0(x)$ считать мало отличающимися или близкими.

Кривые $y = y(x)$ и $y = y_0(x)$ называют близкими в смысле близости нулевого порядка, если модуль их разности $|y(x) - y_0(x)|$ мал.

Кривые $y = y(x)$ и $y = y_0(x)$ называют близкими в смысле близости k-го порядка, если модули разности $\left|y^{(n)}(x) - y_0^{(n)}(x)\right|$ малы $\forall n \in \{0,\ 1,\ 2, \dots ,k\}$.

Тогда в случае близости кривых нулевого порядка можем сказать, что функционал $J[y(x)]$ непрерывен при $y = y_0(x)$, если $\forall \varepsilon > 0\ \ \exists \delta = \delta_\varepsilon > 0$: $\forall y \in M$: $|y(x) - y_0(x)| < \delta$ ($x \in [a,\ b]$ и $x - \forall$) $\to$ $|J[y(x)] - J[y_0(x)]| < \varepsilon$.

Высказывание $|y(x) - y_0(x)| < \delta \quad \forall x \in [a,\ b]$ удобно переписать в виде: $\max\limits_{x \in [a,\ b]} |y(x) - y_0(x)| < \delta$.

И для определения непрерывного функционала имеем:
$\forall \varepsilon > 0\ \ \exists \delta = \delta_\varepsilon > 0$: $\forall y \in M$: $\max\limits_{x \in [a,\ b]} |y(x) - y_0(x)| < \delta \ \to\ |J[y(x)] - J[y_0(x)]| < \varepsilon$.

Для функционала из примера 3) $\delta = \dfrac{\varepsilon}{b - a}$.

Для функционала из примера 4) дело обстоит несколько сложнее. Здесь на первый план выступает изменение производной функции.

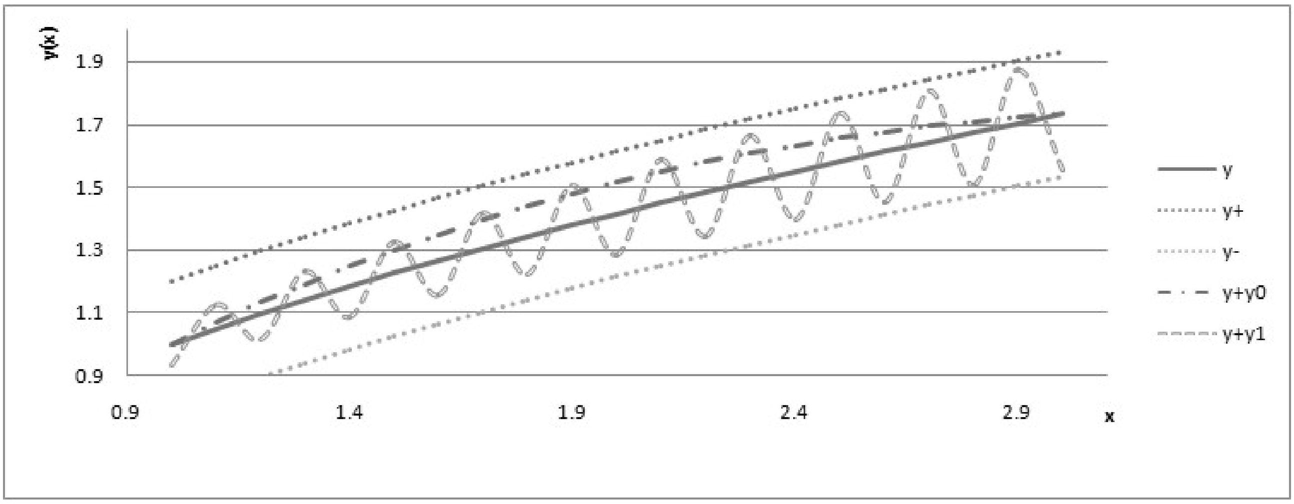

Рис. 1

На рис.1 изображены кривые $y = y(x) + y0$ и $y = y(x) + y1$ близкие в смысле нулевого порядка к кривой $y = y(x)$, но кривые $y = y(x) + y1$ и $y = y(x)$ не являются близкими в смысле близости первого порядка.

В этом случае разумно ввести на малое изменение функции следующее ограничение: $\max\limits_{x \in [a,\ b]} |y(x) - y_0(x)| + \max\limits_{x \in [a,\ b]} |y'(x) - y_0'(x)| < \delta$.

И в случае близости кривых первого порядка функционал будем считать непрерывным, если

$$\forall \varepsilon > 0 \quad \exists \delta = \delta_\varepsilon > 0: \quad \forall y \in M: \quad \max_{x \in [a,\ b]} |y(x) - y_0(x)| + \max_{x \in [a,\ b]} |y'(x) - y_0'(x)| < \delta \quad \to$$

$$|J[y(x)] - J[y_0(x)]| < \varepsilon.$$

Заметим, что если выполнено условие близости первого порядка $\max\limits_{x\in[a,\ b]}\left|y(x)-y_0(x)\right|+\max\limits_{x\in[a,\ b]}\left|y'(x)-y_0'(x)\right|<\delta$, то кривые близки и в смысле близости нулевого порядка $\max\limits_{x\in[a,\ b]}\left|y(x)-y_0(x)\right|<\delta$.

Обратное неверно: для функций $y_0=x$ и $y=x+\delta\sin\dfrac{x}{\delta}$ имеет место $\max\limits_{x\in[a,\ b]}\left|y(x)-y_0(x)\right|=\max\limits_{x\in[a,\ b]}\left|\delta\sin\dfrac{x}{\delta}\right|<\delta$, но $\max\limits_{x\in[a,\ b]}\left|y'(x)-y_0'(x)\right|=\max\limits_{x\in[a,\ b]}\left|\cos\dfrac{x}{\delta}\right|=1$, если $b-a>\delta\cdot 2\pi$.

В дальнейшем представляются важными следующие линейные нормированные пространства:

1. $C([a,\ b])$ – пространство всех непрерывных на $[a,\ b]$ функций с нормой $\left\|y(x)\right\|_C=\max\limits_{x\in[a,\ b]}\left|y(x)\right|$.
2. $C^1([a,\ b])$ – пространство всех непрерывно дифференцируемых на $[a,\ b]$ функций с нормой $\left\|y(x)\right\|_{C^1}=\max\limits_{x\in[a,\ b]}\left|y(x)\right|+\max\limits_{x\in[a,\ b]}\left|y'(x)\right|$.
3. $C^k([a,\ b])$ – пространство всех k раз непрерывно дифференцируемых на $[a,\ b]$ функций с нормой $\left\|y(x)\right\|_{C^k}=\sum\limits_{n=0}^{k}\max\limits_{x\in[a,\ b]}\left|y^{(n)}(x)\right|$.

Под классом M функций, фигурирующих в определении функционала в дальнейшем будем понимать $M=\left\{y(x):\ y(x)\in C^k([a,\ b])\right\}=M_k([a,\ b]),\ k=0,\ 1,\ 2,\ \ldots$.

Теперь можно сформулировать

Определение. Функционал $J[y(x)]$, где $y(x)\in M_k([a,\ b])$ называется ***непрерывным*** при $y_0(x)\in M_k([a,\ b])$, если $\forall\varepsilon>0\ \ \exists\delta=\delta_\varepsilon>0:\ \forall y(x)\in M_k([a,\ b])$: $\left\|y(x)-y_0(x)\right\|_{C^k}<\delta\ \to\ \left|J[y(x)]-J[y_0(x)]\right|<\varepsilon$.

Важным частным случаем функционалов являются линейные функционалы.

Определение. Функционал $J[y(x)]$ называется линейным, если он

1. непрерывен,
2. $\forall\ y_1(x),\ y_2(x)\in M$ и $\forall\ \alpha_1,\ \alpha_2\in\mathbb{R}$ $J[\alpha_1y_1(x)+\alpha_2y_2(x)]=$ $=\alpha_1J[y_1(x)]+\alpha_2J[y_2(x)]$.

Согласно этому определению функционалы в примерах 1) и 3) линейные, в примерах 2) и 4) не линейные.

Перейдем теперь к понятию дифференциала функции.

Напомним, что из курса математического анализа известно, что если приращение функции $\Delta f=f(x+\Delta x)-f(x)$ может быть представлено в виде

$\Delta f = A(x)\Delta x + \alpha(x, \Delta x)\Delta x$, где $A(x)$ не зависит от Δx, а $\alpha(x, \Delta x)\xrightarrow[\Delta x\to 0]{}0$, то функция $f(x)$ называется *дифференцируемой*, а линейная по отношению к Δx часть приращения функции $A(x)\Delta x$ называется *дифференциалом функции* и обозначается: $df = A(x)\Delta x$, $\Delta f = df + \alpha(x, \Delta x)\Delta x$. Производная функции $f'(x) = \lim\limits_{\Delta x\to 0}\dfrac{\Delta f}{\Delta x} = A(x)$, т.е. $df = f'(x)\Delta x$.

Рассмотрим приращение функционала
$\Delta J = J[y(x)+h(x)] - J[y(x)]$,
отвечающее приращению $h(x)\in M([a, b])$ "независимой переменной" $y(x)\in M([a, b])$.

Если приращение функционала можно представить в виде
$\Delta J = L[y(x), h(x)] + \beta[y(x), h(x)]\max\limits_{x\in[a, b]}|h(x)|$,
где $L[y(x), h(x)]$ – линейный по отношению к $h(x)$ функционал, а $\beta[y(x), h(x)]\xrightarrow[\max\limits_{x\in[a,b]}|h(x)|\to 0]{}0$, то $L[y(x), h(x)]$ *называется вариацией функционала* (*по Фреше*) и обозначается $\delta J_y[h] = L[y(x), h(x)]$.

Подытожим вышесказанное, дав определение *сильного дифференциала – дифференциала Фреше*.

Определение. Функционал $J[y(x)]$ называется дифференцируемым при $y = y_0(x)\in M$, если существует линейный по отношению к $h(x)\in M$ функционал $L[y_0(x), h(x)]$ такой, что
$\Delta J = J[y_0(x)+h(x)] - J[y_0(x)] = L[y_0, h] + o(\|h\|_C)$
при $\|h\|_C \to 0$.
Линейная часть приращения $L[y_0(x), h(x)]$ *называется* ***первой вариацией (Фреше) – первым дифференциалом Фреше*** – и обозначается как
$\delta J_{y_0}[h] = L[y_0(x), h(x)]$.

Итак, вариация функционала – это главная, линейная по отношению к приращению функции $h(x)$ часть приращения функционала.

При исследовании функционалов вариация играет ту же роль, что и дифференциал для функции.

Можно дать другое, почти эквивалентное, определение вариации функционала в смысле производной по параметру.

Определение. ***Первой вариацией*** (***по Гато***) функционала при $y = y_0(x)\in M$, если $h(x)\in M$, называется производная от функционала $J[y_0(x) + t\cdot h(x)]$ по параметру t при $t = 0$:

$$\left.\sigma J_{y_0}[h] = \frac{d}{dt} J\left[y_0(x) + t \cdot h(x)\right]\right|_{t=0}.$$

Это определение несколько шире первого, так как существуют примеры функционалов, из приращения которых нельзя выделить главной линейной части, но вариации в смысле определения Гато существуют.

Если провести аналогию с функцией двух независимых переменных, то вариации Фреше можно сопоставить дифференцируемость функции в точке, а вариации Гато – производную по направлению. Существование производных по любому направлению недостаточно для существования дифференциала функции.

Пример 1.1. $f(x, y) = \sqrt{2xy}$.

□ В полярных координатах $x = \rho\cos\varphi$, $y = \rho\sin\varphi$ имеем $f(x, y) = \rho\sqrt{\sin 2\varphi}$.

Рассмотрим производную функции f по направлению $\mathbf{l}$, составляющему угол α с осью абсцисс в точке $(0, 0)$: $\dfrac{df(0, 0)}{d\mathbf{l}} = \sqrt{\sin 2\alpha}$.

Найдем теперь приращение функции f в точке $(0, 0)$:

$\Delta f(0, 0) = f(\Delta x, \Delta y) - f(0, 0) = \sqrt{2\Delta x\Delta y}$.

И вычислим частные производные в точке $(0, 0)$. Поскольку $f(x, 0) = 0$ и $f(0, y) = 0$, то $\dfrac{\partial f(0, 0)}{\partial x} = 0$ и $\dfrac{\partial f(0, 0)}{\partial y} = 0$.

Если функция f дифференцируема в точке $(0, 0)$, то ее приращение имеет вид

$$\Delta f(0, 0) = \frac{\partial f(0, 0)}{\partial x}\Delta x + \frac{\partial f(0, 0)}{\partial y}\Delta y + o\left(\sqrt{\Delta x^2 + \Delta y^2}\right),\ (\Delta x, \Delta y) \to (0, 0),$$

т.е. $\sqrt{2\Delta x\Delta y} = o\left(\sqrt{\Delta x^2 + \Delta y^2}\right)$, $(\Delta x, \Delta y) \to (0, 0)$.

Но это не так, поскольку $\lim\limits_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \dfrac{\sqrt{2\Delta x\Delta y}}{\sqrt{\Delta x^2 + \Delta y^2}} \neq 0$. Действительно, при $\Delta x = \Delta y \to 0$ имеем $\dfrac{\sqrt{2\Delta x\Delta y}}{\sqrt{\Delta x^2 + \Delta y^2}} = \dfrac{\sqrt{2\Delta x^2}}{\sqrt{2\Delta x^2}} = 1$.

Полученное противоречие показывает ошибочность предположения о дифференцируемости функции f в точке $(0, 0)$.■

Замечание 1.1. В прикладном плане определение вариации по Гато (в смысле производной по параметру при начальном значении параметра) более удобно при выводе ряда соотношений вариационного исчисления.

Замечание 1.2. Во втором случае $h(x)$ не является малой величиной, а задает "форму приращения". Малость приращения регулируется параметром t, стремящимся к нулю.

Лемма 1.1. Если функционал $J[y(x)]$, $y(x) \in M$, дифференцируем (по Фреше) при

$y = y_0(x)$, то его вариация $\forall\ h(x) \in M$ равна производной по параметру при начальном значении параметра (вариации по Гато) $\delta J_{y_0}[h] = \frac{d}{dt} J[y_0(x) + t \cdot h(x)]\Big|_{t=0}$.

○ Так как $h(x) \in M$, то $t \cdot h(x) \in M$. $y_0(x) \in M$ поэтому $(y_0(x) + t \cdot h(x)) \in M$.
Функционал дифференцируем, т.е.
$\Delta J = J[y(x) + t \cdot h(x)] - J[y(x)] = \delta J_{y_0}[t \cdot h] + o(\|t \cdot h\|_C)$,
где $\delta J_{y_0}[t \cdot h]$ – линейная часть,
т.е $\delta J_{y_0}[t \cdot h] = t \cdot \delta J_{y_0}[h]$, и
$\Delta J = t \cdot \delta J_{y_0}[h] + o(\|t \cdot h\|_C)$.
Производная от $J[y(x) + t \cdot h(x)]$ по t при $t = 0$ равна

$$\sigma J_{y_0}[h] = \frac{d}{dt} J[y_0(x) + t \cdot h(x)]\Big|_{t=0} = \lim_{t \to 0} \frac{\Delta J}{t} = \delta J_{y_0}[h] + \lim_{t \to 0} \frac{o(\|t \cdot h\|_C)}{t} = \delta J_{y_0}[h]. \ \bullet$$

Итак: если существует вариация в смысле главной линейной части функционала, то существует и вариация в смысле производной по параметру при начальном значении параметра. И в этом случае оба определения "эквивалентны".

Методы нахождения наибольших и наименьших значений функционалов суть вариационного исчисления.

Определение. Функционал $J[y(x)]$ достигает на кривой $y = y_0(x) \in M_k$ ***локального максимума*** (***минимума***), если $\exists \delta > 0 :$ $\forall\ y(x) \in M_k :$ $\|y(x) - y_0(x)\|_{C^k} < \delta$ выполняется неравенство $J[y(x)] \le (\ge) J[y_0(x)]$.

Оба понятия (локальный максимум и локальный минимум) объединяются единым понятием ***локальный экстремум***.

Замечание 1.3.1. Если локальный экстремум достигается на кривых, близких к $y_0(x)$ в смысле сходимости по норме $C([a,\ b])$ (т.е. $k = 0$), то он называется ***сильным*** локальным экстремумом.

Замечание 1.3.2. Если локальный экстремум достигается лишь в смысле сходимости по норме $C^k([a,\ b])$, $k \ge 1$, то он называется ***слабым*** локальным экстремумом.

Замечание 1.3.3. Если достигается сильный экстремум, то достигается и слабый. Действительно, если $\|y(x) - y_0(x)\|_{C^k} < \delta$, то и $\|y(x) - y_0(x)\|_C < \delta$. Обратное, вообще говоря, не верно: кривая, доставляющая экстремум в некоторой совокупности кривых сравнения, может перестать его давать после расширения этой совокупности кривых. Поэтому кривая, доставляющая слабый экстремум, может не давать сильного.

Замечание 1.3.4 В классическом вариационном исчислении употребляется еще термин ***абсолютного*** (или ***глобального***) экстремума: говорят, что функционал $J[y(x)]$ достигает на кривой $y = y_0(x) \in M_k$ ***абсолютного максимума*** (***минимума***), если $\forall\ y(x) \in M_k$ выполняется неравенство $J[y(x)] \le (\ge) J[y_0(x)]$.

Теорема 1.1. Если дифференцируемый (т.е. имеющий вариацию по Фреше) функционал $J[y(x)]$ достигает экстремума при $y = y_0(x) \in M$, то необходимо, чтобы его первая вариация обращалась в нуль при $y = y_0(x)$: $\delta J_{y_0}[h(x)] = 0$ для $\forall\ h(x)$: $y = y_0(x) \in M$.

○ При конкретных функциях $y_0(x)$ и $h(x)$: $y = y_0(x) \in M$ $J[y_0(x) + t \cdot h(x)] = g(t)$ – функция одной переменной, которая при $t = 0$ достигает экстремума, дифференцируемая в силу условий теоремы,

поэтому необходимо $\left.\dfrac{\partial}{\partial t} g(t)\right|_{t=0} = 0$.

Но $\left.\dfrac{\partial}{\partial t} g(t)\right|_{t=0} = \left.\dfrac{\partial}{\partial t} J[y_0(x) + t \cdot h(x)]\right|_{t=0}$ по построению,

а по **лемме 1.1** $\left.\dfrac{\partial}{\partial t} J[y_0(x) + t \cdot h(x)]\right|_{t=0} = \delta J_{y_0}[h(x)]$,

откуда и получаем $\delta J_{y_0}[h(x)] = 0$. ●

Лемма 1.2. (**основная лемма вариационного исчисления**)

Если $\forall\ \eta(x) \in C^1([a,\ b])$ $\int\limits_a^b \Phi(x)\eta(x)dx = 0$,

где $\Phi(x) \in C([a,\ b])$,

то $\Phi(x) \equiv 0$.

○ Докажем от противного:

пусть $\Phi(x) \not\equiv 0$ на промежутке $[a,\ b]$.

В силу непрерывности $\Phi(x)$ $\exists \hat{x} \in [a,\ b]$: $\Phi(\hat{x}) \neq 0$, например $\Phi(\hat{x}) > 0$ (случай $\Phi(\hat{x}) < 0$ рассматривается аналогично).

Тогда, как известно из курса математического анализа, $\exists \alpha,\ \beta \in [a,\ b]$: $\hat{x} \in [\alpha,\ \beta]$ и $\forall x \in [\alpha,\ \beta]$ $\Phi(x) > 0$.

Функция $\eta(x)$ любая из класса $C^1([a,\ b])$, поэтому выберем $\eta(x) = \hat{\eta}(x)$ так, чтобы

- $\hat{\eta}(x) \ge 0$ $\forall x \in [\alpha,\ \beta]$,
- $\hat{\eta}(x) \not\equiv 0$ $\forall x \in [\alpha,\ \beta]$,
- $\hat{\eta}(x) \equiv 0$ $\forall x \notin [\alpha,\ \beta]$.

Например,

при $a < \alpha < \beta < b$

$$\hat{\eta}(x) = \begin{cases} (x-\alpha)^2 (x-\beta)^2,\ x \in [\alpha,\ \beta], \\ 0,\ x \in [a,\ \alpha] \cup [\beta,\ b], \end{cases}$$

при $a = \alpha < \beta < b$ $\qquad \hat{\eta}(x) = \begin{cases} (x-a)^2(x-\beta)^2, & x \in [a, \beta], \\ 0, & x \in [\beta, b], \end{cases}$

при $a < \alpha < \beta = b$ $\qquad \hat{\eta}(x) = \begin{cases} (x-\alpha)^2(x-b)^2, & x \in [\alpha, b], \\ 0, & x \in [a, \alpha]. \end{cases}$

Тогда $\int\limits_a^b \Phi(x)\eta(x)dx > 0$, т.к. подынтегральная функция положительна. при этом все условия леммы выполнены.

Полученное противоречие доказывает лемму. ●

Определение. Функции $y(x) \in M$, на которых задан функционал $J[y(x)]$, будем называть ***допустимыми***.

Определение. Пусть $y_0(x) \in M$. Приращение $h(x)$ функции $y_0(x)$ назовем **допустимым**, если $y_0(x) + h(x) \in M$.

§2. Простейшая задача вариационного исчисления.

В приложениях часто встречаются функционалы вида:

$$J[y(x)] = \int\limits_a^b F(x, y(x), y'(x))dx, \qquad (2.1)$$

где $F(x, y, p)$ – заданная дважды непрерывно дифференцируемая функция для $\forall x \in [a,b]$ и $\forall (y, p) \in R^2_{(y,p)}$ – плоскости с декартовыми прямоугольными координатами y, p.

Простейшей вариационной задачей (ПВЗ) называется задача нахождения слабого локального экстремума функционала (2.1) в классе

$$M = \{y(x):\ y(x) \in C^1([a, b]),\ y(a) = A,\ y(b) = B\} \qquad (2.2)$$

– непрерывно дифференцируемых на $[a,b]$ функций $y(x)$, удовлетворяющих граничным условиям с фиксированным значением допустимой функции на концах промежутка – ***задача с закрепленными концами***.

Обозначим за $C^{*1}([a, b])$ пространство функций $f(x) \in C^1([a, b])$, на которые наложено дополнительное ограничение в виде однородных граничных условий: $f(a) = f(b) = 0$.

Допустимыми приращениями для ПВЗ будут

$$h(x) \in M^* = \{h(x):\ h(x) \in C^{*1}([a, b])\} \qquad (2.3)$$

Лемма 2.1. Пусть для $\Phi(x) \in C([a, b])$ и $\forall\ \eta(x) \in C^{*1}([a, b])$ имеет место

$$\int\limits_a^b \Phi(x)\eta'(x)dx = 0,$$

тогда $\Phi(x) \equiv const$ для $\forall\ x \in [a,b]$.

○ Обозначим $\eta'(x) = k(x)$. Т.к. $\forall\ \eta(x) \in C^{*^1}([a,\, b])$, то $k(x) \in C([a,\, b])$.

Тогда $\eta(x) = \int\limits_a^x k(\xi)d\xi$,

причем $\eta(a) = 0$ по построению,

а $\eta(b) = \int\limits_a^b k(\xi)d\xi = 0$ по условию.

$\int\limits_a^b \Phi(\xi)d\xi = A$. Положим $C = \dfrac{A}{b-a}$. Тогда $\int\limits_a^b (\Phi(\xi) - C)d\xi = 0$.

Рассмотрим функцию $g(x) = \int\limits_a^x (\Phi(\xi) - C)d\xi \in M^*$.

Так как $\eta(x) \in C^{*^1}([a,\, b])$ – произвольная, то можем выбрать $\eta(x) = g(x)$, при этом $\eta'(x) = k(x) = \Phi(x) - C$.

Рассмотрим теперь интеграл $I = \int\limits_a^b (\Phi(\xi) - C)^2 d\xi$.

Т.к. $(\Phi(\xi) - C)^2 \geq 0$, то и $\underline{I \geq 0}$.

С другой стороны $I = \int\limits_a^b \Phi(\xi)(\Phi(\xi) - C)d\xi - C\int\limits_a^b (\Phi(\xi) - C)d\xi$.

- При этом $\int\limits_a^b \Phi(\xi)(\Phi(\xi) - C)dx = \int\limits_a^b \Phi(\xi)k(\xi)d\xi = \int\limits_a^b \Phi(\xi)\eta'(\xi)d\xi = 0$ по условию,
- а $\int\limits_a^b (\Phi(\xi) - C)d\xi = 0$ по построению,

т.е. $I = 0$.

Но тогда $(\Phi(x) - C)^2 = 0$ для $\forall\ x \in [a, b]$,
т.е. $\Phi(x) = C = const$ для $\forall\ x \in [a, b]$. ●