Ekstremalne zastosowania matematyki w fizyce, czyli o rachunku wariacyjnym słów kilka

Sebastian Haratyk

9 maja 2015

Definicja 1

Odwzorowanie $J : \mathfrak{F} \to \mathbb{R}$, gdzie $\mathfrak{F}$ jest pewną rodziną funkcji, nazywamy funkcjonałem.

Definicja 1

Odwzorowanie $J : \mathfrak{F} \to \mathbb{R}$, gdzie $\mathfrak{F}$ jest pewną rodziną funkcji, nazywamy funkcjonałem.

Przykład 1. Długość krzywej y łączącej dwa ustalone punkty (x_1, y_1), (x_2, y_2), przy założeniu, że jest ona klasy C^1 na przedziale $[x_1, x_2]$, jest przykładem funkcjonału. Jak wiadomo, można ją obliczyć z następującego wzoru:

$$J(y) = \int\limits_{x_1}^{x_2} \sqrt{1 + y'(x)^2} dx \, .$$

Niech $a, b, A, B \in \mathbb{R}$ oraz $a < b$.
W naszych dalszych rozważaniach, będziemy rozpatrywać przestrzenie metryczne zupełne następujących postaci :

(i) $(C([a,b]), \rho_0)$, gdzie $C([a,b])$ jest przestrzenią funkcji ciągłych na przedziale $[a,b]$, a $\rho_0 : C([a,b]) \times C([a,b]) \to \mathbb{R}$ jest metryką zdefiniowaną wzorem:

$$\rho_0(u,v) = \max_{x\in[a,b]} |v(x) - u(x)| \qquad \text{dla } u, v \in C([a,b]);$$

Niech $a, b, A, B \in \mathbb{R}$ oraz $a < b$.
W naszych dalszych rozważaniach, będziemy rozpatrywać przestrzenie metryczne zupełne następujących postaci :

(i) $(C([a,b]), \rho_0)$, gdzie $C([a,b])$ jest przestrzenią funkcji ciągłych na przedziale $[a,b]$, a $\rho_0 : C([a,b]) \times C([a,b]) \to \mathbb{R}$ jest metryką zdefiniowaną wzorem:

$$\rho_0(u,v) = \max_{x\in[a,b]} |v(x) - u(x)| \qquad \text{dla } u, v \in C([a,b]);$$

(ii) $(C(a,b,A,B), \hat{\rho}_0)$, gdzie $C(a,b,A,B) = \{y \in C([a,b]) : y(a) = A \wedge y(b) = B\}$ (tzw. przestrzeń funkcji umocowanych na końcach), a $\hat{\rho}_0 = \rho_0 \upharpoonright_{C(a,b,A,B)\times C(a,b,A,B)}$;

(iii) $(C^1([a,b]),\rho_1)$, gdzie $C^1([a,b])$ oznacza przestrzeń funkcji różniczkowalnych o ciągłych pochodnych pierwszego rzędu (w przedziale $[a,b]$), natomiast
$\rho_1 : C^1([a,b]) \times C^1([a,b]) \to \mathbb{R}$ jest metryką określoną wzorem:

$$\rho_1(u,v) = \max_{x\in[a,b]} |v(x)-u(x)| + \max_{x\in[a,b]} |v'(x)-u'(x)| = \rho_0(u,v)+\rho_0(u',v'$$

dla $u, v \in C^1([a,b])$;

(iii) $(C^1([a,b]), \rho_1)$, gdzie $C^1([a,b])$ oznacza przestrzeń funkcji różniczkowalnych o ciągłych pochodnych pierwszego rzędu (w przedziale $[a,b]$), natomiast $\rho_1 : C^1([a,b]) \times C^1([a,b]) \to \mathbb{R}$ jest metryką określoną wzorem:

$$\rho_1(u,v) = \max_{x\in[a,b]} |v(x) - u(x)| + \max_{x\in[a,b]} |v'(x) - u'(x)| = \rho_0(u,v) + \rho_0(u',v'$$

dla $u, v \in C^1([a,b])$;

(iv) $(C^1(a,b,A,B), \hat{\rho}_1)$, gdzie $\hat{\rho}_1 = \rho_1 \restriction_{C^1(a,b,A,B)\times C^1(a,b,A,B)}$ oraz $C^1(a,b,A,B) = \{y \in C^1([a,b]) : y(a) = A \wedge y(b) = B\}$.

Dla ułatwienia dalszych rozważań, przyjmiemy oznaczeniowo, iż $\hat{\rho}_0 = \rho_0$ oraz $\hat{\rho}_1 = \rho_1$. Ponado oznaczmy przez $K_0(x, r)$ kulę o środku w punkcie (funkcji) x i promieniu r w przestrzeniach $(C([a, b]), \rho_0)$ i $(C(a, b, A, B), \hat{\rho}_0)$, natomiast w pozostałych dwóch, z wspomnianych wcześniej przestrzeni, kulę będziemy oznaczać jako $K_1(x, r)$.

Stwierdzenie 1

Dla dowolnej funkcji $u \in C^1([a,b])$ ($u \in C^1(a,b,A,B)$) oraz dowolnego $r > 0$ zachodzi
$$K_1(u,r) \subset K_0(u,r)$$

Stwierdzenie 1

Dla dowolnej funkcji $u \in C^1([a, b])$ ($u \in C^1(a, b, A, B)$) oraz dowolnego $r > 0$ zachodzi

$$K_1(u, r) \subset K_0(u, r)$$

Dowód. Ustalmy $v \in K_1(u, r)$. Wówczas

$$r > \rho_1(u, v) = \rho_0(u, v) + \rho(u', v')$$

Korzystając z nieujemności metryki wnosimy, iż $\rho_0(u, v) < r$, a zatem $v \in K_0(u, r)$.

Definicja 2

Niech $J : X \to \mathbb{R}$ będzie funkcjonałem określonym na przestrzeni metrycznej X oraz $y \in X$.
Mówimy, że w punkcie y osiagane jest minimum (maksimum) lokalne funkcjonału J, jeżeli istnieje taka kula $K(y, r) \subset X$, że $J(v) \geq J(y)$ $(J(v) \leq J(y))$ dla $v \in K(y, r)$.

Definicja 2

Niech $J : X \to \mathbb{R}$ będzie funkcjonałem określonym na przestrzeni metrycznej X oraz $y \in X$.
Mówimy, że w punkcie y osiagane jest minimum (maksimum) lokalne funkcjonału J, jeżeli istnieje taka kula $K(y, r) \subset X$, że $J(v) \geq J(y)$ $(J(v) \leq J(y))$ dla $v \in K(y, r)$. Punkt y nazywamy ekstremum lokalnym.
Ekstrema w przestrzeniach $C([a, b])$ oraz $C(a, b, A, B)$ nazywamy mocnymi, natomiast w przypadku przestrzeni $C^1([a, b])$, $C^1(a, b, A, B)$ ekstrema lokalne nazywamy słabymi.

Twierdzenie 1

Jeżeli na pewnej funkcji klasy C^1 osiągane jest minimun (maksimum) mocne, to na tej funkcji osiągne jest minimum (maksimum) słabe.

Twierdzenie 1

Jeżeli na pewnej funkcji klasy C^1 osiągane jest minimun (maksimum) mocne, to na tej funkcji osiągne jest minimum (maksimum) słabe.

Dowód. Udowodnimy jedynie przypadek, gdy funkcja y (klasy C^1) jest maksimum mocnym funkcjonału J w $C(a, b, A, B)$. Wówczas istnieje taka kula $K_0(y, r)$ w tej przestrzeni, że dla każdego $v \in K_0(y, r)$ zachodzi $J(y) \geq J(v)$. Na mocy stwierdzenia 1, dla każdego $x \in K_1(y, r)$ mamy, iż $x \in K_0(y, r)$. Stąd

$$\bigwedge_{x \in K_1(y,r)} J(y) \geq J(x),$$

zatem na funkcji y osiągane jest maksimum słabe.

Twierdzenie 1

Jeżeli na pewnej funkcji klasy C^1 osiągane jest minimun (maksimum) mocne, to na tej funkcji osiągne jest minimum (maksimum) słabe.

Dowód. Udowodnimy jedynie przypadek, gdy funkcja y (klasy C^1) jest maksimum mocnym funkcjonału J w $C(a, b, A, B)$. Wówczas istnieje taka kula $K_0(y, r)$ w tej przestrzeni, że dla każdego $v \in K_0(y, r)$ zachodzi $J(y) \geq J(v)$. Na mocy stwierdzenia 1, dla każdego $x \in K_1(y, r)$ mamy, iż $x \in K_0(y, r)$. Stąd

$$\bigwedge_{x \in K_1(y,r)} J(y) \geq J(x),$$

zatem na funkcji y osiągane jest maksimum słabe.

Twierdzenie odwrotne do powyższego twierdzenia nie zachodzi.

Twierdzenie 2

Niech $f : [a, b] \times [c, d] \to \mathbb{R}$ będzie ciągła funkcją o ciągłej pochodnej $\frac{\partial f}{\partial x}$ w $[a, b] \times [c, d]$.
Wówczas dla każdego $x \in [a, b]$

$$\frac{d}{dx} \int_c^d f(x, y) dy = \int_c^d \frac{\partial}{\partial x} f(x, y) dy.$$

Twierdzenie 3 (Cauchy'ego-Buniakowskiego AKA Schwarza dla całek)

Jeżeli funkcje f, g są ciągłe na $[a, b]$, to

$$\left(\int_a^b f(x)g(x)dx\right)^2 \leq \int_a^b f(x)^2 dx \int_a^b g(x)^2 dx.$$

Twierdzenie 3 (Cauchy'ego-Buniakowskiego AKA Schwarza dla całek)

Jeżeli funkcje f, g są ciągłe na $[a, b]$, to

$$\left(\int_a^b f(x)g(x)dx\right)^2 \leq \int_a^b f(x)^2 dx \int_a^b g(x)^2 dx.$$

Lemat 4

Niech $f, g \in C^1([a, b])$. Jeżeli dla każdej funkcji $h \in C^1(a, b, 0, 0)$

$$\int_a^b (f(x)h(x) + g(x)h'(x))dx = 0,$$

to funkcja $g \in C^1([a, b])$ oraz $g' = f$.

Twierdzenie 5 (warunek konieczny istnienia ekstremum)

Niech $F : [a, b] \times \mathbb{R}^2 \to \mathbb{R}$ będzie funkcją klasy C^1 oraz J będzie funkcjonałem w przestrzeni $C^1(a, b, A, B)$ określonym wzorem:

$$J(u) = \int_a^b F(x, u(x), u'(x))dx, \qquad u \in C^1(a, b, A, B)$$

Jeżeli na funkcji $y \in C^1(a, b, A, B)$ osiągne jest słabe ekstremum funkcjonału J, to dla $x \in [a, b]$ funkcja $\frac{\partial}{\partial y'} F(x, y(x), y'(x))$ jest różniczkowalna oraz spełniona jest równość

$$\frac{d}{dx}\left[\frac{\partial F}{\partial y'}(x, y(x), y'(x))\right] - \frac{\partial F}{\partial y}(x, y(x), y'(x)) = 0.$$

Dowód. Załóżmy, że funkcjonał J osiąga słabe ekstremum na funkcji $y \in C^1(a, b, A, B)$, tzn. istnieje taka kula $K_1(y, r)$, że dla każdego $u \in K_1(y, r)$ zachodzi $J(u) \geq J(y)$ (gdy na y osiągane jest minimum) albo $J(u) \leq J(y)$ (dla maksimum).

Dowód. Załóżmy, że funkcjonał J osiąga słabe ekstremum na funkcji $y \in C^1(a, b, A, B)$, tzn. istnieje taka kula $K_1(y, r)$, że dla każdego $u \in K_1(y, r)$ zachodzi $J(u) \geq J(y)$ (gdy na y osiągane jest minimum) albo $J(u) \leq J(y)$ (dla maksimum).
Weźmy teraz dowolną funkcję $h \in C^1(a, b, 0, 0)$. Jako funkcja klasy C^1 na zbiorze zwartym $[a, b]$ jest ona wraz z swoją pochodną ograniczona, tzn.

$$\bigvee_{k>0} \max_{x\in[a,b]} |h(x)| + \max_{x\in[a,b]} |h'(x)| < k.$$

Dowód. Załóżmy, że funkcjonał J osiąga słabe ekstremum na funkcji $y \in C^1(a, b, A, B)$, tzn. istnieje taka kula $K_1(y, r)$, że dla każdego $u \in K_1(y, r)$ zachodzi $J(u) \geq J(y)$ (gdy na y osiągane jest minimum) albo $J(u) \leq J(y)$ (dla maksimum).
Weźmy teraz dowolną funkcję $h \in C^1(a, b, 0, 0)$. Jako funkcja klasy C^1 na zbiorze zwartym $[a, b]$ jest ona wraz z swoją pochodną ograniczona, tzn.

$$\bigvee_{k>0} \max_{x \in [a,b]} |h(x)| + \max_{x \in [a,b]} |h'(x)| < k.$$

Niech $\delta_0 = \frac{r}{k}$. Rozważmy dla $\delta \in \mathbb{R}$ takich, że $|\delta| < \delta_0$ rodzinę funkcji $u_\delta = y + \delta h$. Wykażemy, że $u_\delta \in K_1(y, r)$.

$$u_\delta = y + \delta h$$

Wykażemy, że $u_\delta \in K_1(y, r)$.

$$u_\delta = y + \delta h$$

Wykażemy, że $u_\delta \in K_1(y, r)$.
Ponieważ $y, h \in C^1([a, b])$, więc również $u_\delta \in C^1([a, b])$. Z faktów, iż $y(a) = A$, $y(b) = B$ i $h(a) = h(b) = 0$ wynikają równości $u_\delta(a) = A$ oraz $u_\delta(b) = B$, a zatem funkcje $u_\delta \in C^1(a, b, A, B)$.

$$u_\delta = y + \delta h$$

Wykażemy, że $u_\delta \in K_1(y, r)$.
Ponieważ $y, h \in C^1([a, b])$, więc również $u_\delta \in C^1([a, b])$. Z faktów, iż $y(a) = A$, $y(b) = B$ i $h(a) = h(b) = 0$ wynikają równości $u_\delta(a) = A$ oraz $u_\delta(b) = B$, a zatem funkcje $u_\delta \in C^1(a, b, A, B)$. Wtedy

$$\rho_1(u_\delta, y) = \max_{x\in[a,b]} |u_\delta(x) - y(x)| + \max_{x\in[a,b]} |u_\delta'(x) - y'(x)| =$$

$$= \max_{x\in[a,b]} |\delta h(x)| + \max_{x\in[a,b]} |\delta h'(x)| =$$

$$= |\delta| \left(\max_{x\in[a,b]} |h(x)| + \max_{x\in[a,b]} |h'(x)| \right) < \delta_0 k = r,$$

a zatem $u_\delta \in K_1(y, r)$ (gdy $|\delta| < \delta_0$).

W związku z poczynionym założeniem mamy, iż $J(u_\delta) \geq J(y)$ albo $J(u_\delta) \leq J(y)$ (w zależności od przypadku ekstremum), co przy oznaczeniu

$$\varphi(\delta) = \int_a^b F(x, y(x) + \delta h(x), y'(x) + \delta h'(x))dx$$

daje $\varphi(\delta) \geq \varphi(0)$ czy też $\varphi(\delta) \leq \varphi(0)$.

[1]Dla formalności, naszą funkcję podcałkową F o trzech zmiennych, możemy interpretować jako funkcję dwóch zmiennych - x oraz δ.

W związku z poczynionym założeniem mamy, iż $J(u_\delta) \geq J(y)$ albo $J(u_\delta) \leq J(y)$ (w zależności od przypadku ekstremum), co przy oznaczeniu

$$\varphi(\delta) = \int_a^b F(x, y(x) + \delta h(x), y'(x) + \delta h'(x))dx$$

daje $\varphi(\delta) \geq \varphi(0)$ czy też $\varphi(\delta) \leq \varphi(0)$. Stąd wnioskujemy, iż funkcja φ określona dla $\delta \in (-\delta_0, \delta_0)$ ma lokalne ekstremum w punkcie $\delta = 0$. Z postaci funckji φ oraz twierdzenia 2 [1] wynika, że jest ona różniczkowalna, a zatem skoro ma ekstremum lokalne dla $\delta = 0$, to $\varphi'(0) = 0$.

[1]Dla formalności, naszą funkcję podcałkową F o trzech zmiennych, możemy interpretować jako funkcję dwóch zmiennych - x oraz δ.

Otrzymujemy więc

$$\varphi'(\delta) = \frac{d}{d\delta}\int\limits_a^b F(x, y(x) + \delta h(x), y'(x) + \delta h'(x))dx =$$

$$= \int\limits_a^b [\frac{\partial}{\partial y} F(x, y(x) + \delta h(x), y'(x) + \delta h'(x))h(x)dx +$$

$$+ \frac{\partial}{\partial y'} F(x, y(x) + \delta h(x), y'(x) + \delta h'(x))h'(x)dx]dx,$$

Otrzymujemy więc

$$\varphi'(\delta) = \frac{d}{d\delta}\int\limits_a^b F(x, y(x) + \delta h(x), y'(x) + \delta h'(x))dx =$$

$$= \int\limits_a^b [\frac{\partial}{\partial y} F(x, y(x) + \delta h(x), y'(x) + \delta h'(x))h(x)dx +$$

$$+ \frac{\partial}{\partial y'} F(x, y(x) + \delta h(x), y'(x) + \delta h'(x))h'(x)dx]dx,$$

a zatem w punkcie $\delta = 0$

$$\int\limits_a^b \left[\frac{\partial}{\partial y} F(x, y(x), y'(x))h(x)dx + \frac{\partial}{\partial y'} F(x, y(x), y'(x))h'(x)dx\right] dx = 0$$

Oznaczmy $f(x) = \frac{\partial}{\partial y} F(x, y(x), y'(x))$ i $g(x) = \frac{\partial}{\partial y'} F(x, y(x), y'(x))$. Zauważmy, że obie te funkcje są ciągłe oraz dla dowolnych funkcji $h \in C^1(a, b, 0, 0)$

$$\int_a^b [f(x)h(x) + g(x)h'(x)]dx = 0.$$

Oznaczmy $f(x) = \frac{\partial}{\partial y} F(x, y(x), y'(x))$ i $g(x) = \frac{\partial}{\partial y'} F(x, y(x), y'(x))$. Zauważmy, że obie te funkcje są ciągłe oraz dla dowolnych funkcji $h \in C^1(a, b, 0, 0)$

$$\int_a^b [f(x)h(x) + g(x)h'(x)]dx = 0.$$

Na mocy lematu 1 otrzymujemy, iż g jest różniczkowalna oraz $g' = f$, a zatem

$$\frac{d}{dx}\left[\frac{\partial}{\partial y'} F(x, y(x), y'(x))\right] = \frac{\partial}{\partial y} F(x, y(x), y'(x)),$$

Oznaczmy $f(x) = \frac{\partial}{\partial y} F(x, y(x), y'(x))$ i $g(x) = \frac{\partial}{\partial y'} F(x, y(x), y'(x))$. Zauważmy, że obie te funkcje są ciągłe oraz dla dowolnych funkcji $h \in C^1(a, b, 0, 0)$

$$\int_a^b [f(x)h(x) + g(x)h'(x)]dx = 0.$$

Na mocy lematu 1 otrzymujemy, iż g jest różniczkowalna oraz $g' = f$, a zatem

$$\frac{d}{dx}\left[\frac{\partial}{\partial y'} F(x, y(x), y'(x))\right] = \frac{\partial}{\partial y} F(x, y(x), y'(x)),$$

skąd ostatecznie otrzymujemy

$$\frac{d}{dx}\left[\frac{\partial}{\partial y'} F(x, y(x), y'(x))\right] - \frac{\partial}{\partial y} F(x, y(x), y'(x)) = 0.$$

■

Definicja 3

Mówimy, że funkcja $y \in C^1(a, b, A, B)$ jest ekstremalą funkcjonału J (określonego wzorem $J(u) = \int\limits_a^b F(x, u(x), u'(x))dx,\ u \in C^1(a, b, A, B)$), gdy dla $x \in [a, b]$ spełnia ona równanie

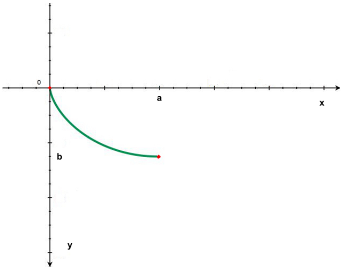
$$\frac{d}{dx}\left[\frac{\partial F}{\partial y'}(x, y, y')\right] - \frac{\partial F}{\partial y}(x, y, y') = 0,$$

zwane równaniem Euler'a.

Zagadnienie brachistochrony jest najstarszym niegeometrycznym zagadnieniem rachunku wariacyjnego postawionym i rozwiązanym przez Jana Bernoulli'ego w 1696 roku.

Zagadnienie brachistochrony jest najstarszym niegeometrycznym zagadnieniem rachunku wariacyjnego postawionym i rozwiązanym przez Jana Bernoulli'ego w 1696 roku.

Rozważamy dwa punkty - początek układu (w którym oś y wskazuje w dół) oraz punkt (a, b) będący punktem położonym niżej (jak na rysunku poniżej). Naszym zadaniem będzie wyznaczenie krzywej, po której punkt materialny przemieści się (wyłącznie pod wpływem siły grawitacji) w jak najkrótszym czasie od punktu $(0, 0)$ do punktu (a, b). Dodatkowo zakładamy, że z początku punkt materialny znajduje się w spoczynku w punkcie $(0, 0)$



Zanim przejdziemy do wyznaczenia wspomnianej krzywej, wykażemy (przy pewnym założeniu) prostszą wersję równania Eulera.

Zanim przejdziemy do wyznaczenia wspomnianej krzywej, wykażemy (przy pewnym założeniu) prostszą wersję równania Eulera.
Przy wspomnianym wyprowadzeniu będziemy stosować zapis uproszczony, np. zamiast $\frac{\partial F}{\partial y'}(x, y, y')$ będziemy pisać po prostu $\frac{\partial F}{\partial y'}$.

Zanim przejdziemy do wyznaczenia wspomnianej krzywej, wykażemy (przy pewnym założeniu) prostszą wersję równania Eulera.
Przy wspomnianym wyprowadzeniu będziemy stosować zapis uproszczony, np. zamiast $\frac{\partial F}{\partial y'}(x, y, y')$ będziemy pisać po prostu $\frac{\partial F}{\partial y'}$.
Załóżmy, że funkcja F nie zależy bezpośrednio od zmiennej niezależnej x.
Zauważamy, iż

$$\frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) = y'\frac{d}{dx}\frac{\partial F}{\partial y'} + y''\frac{\partial F}{\partial y'}.$$

Zanim przejdziemy do wyznaczenia wspomnianej krzywej, wykażemy (przy pewnym założeniu) prostszą wersję równania Eulera.
Przy wspomnianym wyprowadzeniu będziemy stosować zapis uproszczony, np. zamiast $\frac{\partial F}{\partial y'}(x, y, y')$ będziemy pisać po prostu $\frac{\partial F}{\partial y'}$.
Załóżmy, że funkcja F nie zależy bezpośrednio od zmiennej niezależnej x.
Zauważamy, iż

$$\frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) = y'\frac{d}{dx}\frac{\partial F}{\partial y'} + y''\frac{\partial F}{\partial y'}.$$

Korzystając z równania Eulera oraz powyższej równości otrzymujemy

$$\frac{d}{dx}\left(y'\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y''.$$

Różniczka zupełna funkcji F jest postaci:

$$\frac{dF}{dx} = \frac{\partial F}{\partial y} y' + \frac{\partial F}{\partial y'} y'' + \frac{\partial F}{\partial x},$$

ale skoro funkcja F nie zależy jawnie od zmiennej x, to $\frac{\partial F}{\partial x} = 0$, a zatem

$$\frac{d}{dx}\left(y' \frac{\partial F}{\partial y'} \right) = \frac{dF}{dx}.$$

Różniczka zupełna funkcji F jest postaci:

$$\frac{dF}{dx} = \frac{\partial F}{\partial y} y' + \frac{\partial F}{\partial y'} y'' + \frac{\partial F}{\partial x},$$

ale skoro funkcja F nie zależy jawnie od zmiennej x, to $\frac{\partial F}{\partial x} = 0$, a zatem

$$\frac{d}{dx}\left(y' \frac{\partial F}{\partial y'}\right) = \frac{dF}{dx}.$$

Ostatecznie po scałkowaniu powyższej równości otrzymujemy uproszczone równanie Euler'a (przy naszym dodatkowym założeniu)

$$y' \frac{\partial F}{\partial y'} - F = const.$$

Powróćmy zatem do naszego zagadnienia brachistochrony.

Powróćmy zatem do naszego zagadnienia brachistochrony.
Skoro punkt materialny znajduje się początkowo w stanie spoczynku, to z zasady zachowania energii jego energia kinetyczna ($\frac{mv^2}{2}$) będzie równa stracie energii potencjalnej (mgy), czyli $v = \sqrt{2gy}$. Z drugiej strony, wiemy iż $v(t) = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$ jest jego prędkością.

Powróćmy zatem do naszego zagadnienia brachistochrony.
Skoro punkt materialny znajduje się początkowo w stanie spoczynku, to z zasady zachowania energii jego energia kinetyczna ($\frac{mv^2}{2}$) będzie równa stracie energii potencjalnej (mgy), czyli $v = \sqrt{2gy}$. Z drugiej strony, wiemy iż $v(t) = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$ jest jego prędkością. Korzystając z faktu, że jeśli x i y są powiązane parametrem t oraz $\frac{dx}{dt} \neq 0$, to $\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}}$, możemy zapisać równoważnie

$$v(t) = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\frac{dx}{dt}.$$

Uwzględniając prawo zachowania energii oraz korzystając z faktu, iż jeśli $x = f(t)$, to $\frac{dx}{dt} = \frac{1}{\frac{dt}{dx}}$ otrzymujemy

$$\frac{dt}{dx} = \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{\sqrt{2gy}}$$

Uwzględniając prawo zachowania energii oraz korzystając z faktu, iż jeśli $x = f(t)$, to $\frac{dx}{dt} = \frac{1}{\frac{dt}{dx}}$ otrzymujemy

$$\frac{dt}{dx} = \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{\sqrt{2gy}}$$

Stąd wnoismy, iż całkowity czas, w jakim punkt materialny przemieści się po naszej krzywej y od punktu $(0,0)$ do punktu (a,b) wynosi

$$T(y) = \int_0^a t'(x)dx = \int_0^a \frac{\sqrt{1 + y'2}}{\sqrt{2gy}} dx.$$

Uwzględniając prawo zachowania energii oraz korzystając z faktu, iż jeśli $x = f(t)$, to $\frac{dx}{dt} = \frac{1}{\frac{dt}{dx}}$ otrzymujemy

$$\frac{dt}{dx} = \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{\sqrt{2gy}}$$

Stąd wnoismy, iż całkowity czas, w jakim punkt materialny przemieści się po naszej krzywej y od punktu $(0, 0)$ do punktu (a, b) wynosi

$$T(y) = \int_0^a t'(x)dx = \int_0^a \frac{\sqrt{1 + y'2}}{\sqrt{2gy}} dx.$$

W związku z powyższym, będziemy szukać ekstremum funkcjonału zadanego przy pomocy całki z funkcji $F(x, y, y') = \frac{\sqrt{1+y'2}}{\sqrt{2gy}}$, która nie zależy bezpośrednio od x.

Możemy zatem skorzystać z uproszczonej równości Eulera

$$y' \frac{\partial F}{\partial y'} - F = c_1,$$

$$c_1 = \frac{y'^2}{\sqrt{2gy(1+y'^2)}} - \frac{\sqrt{1+y'^2}}{\sqrt{2gy}} = \frac{1}{\sqrt{2gy(1+y'^2)}},$$

Możemy zatem skorzystać z uproszczonej równości Eulera

$$y' \frac{\partial F}{\partial y'} - F = c_1,$$

$$c_1 = \frac{y'^2}{\sqrt{2gy(1+y'^2)}} - \frac{\sqrt{1+y'^2}}{\sqrt{2gy}} = \frac{1}{\sqrt{2gy(1+y'^2)}},$$

co po podniesieniu do kwadratu daje

$$y(1+y'^2) = \frac{1}{2gc_1^2} = c.$$

Możemy zatem skorzystać z uproszczonej równości Eulera

$$y' \frac{\partial F}{\partial y'} - F = c_1,$$

$$c_1 = \frac{y'^2}{\sqrt{2gy(1+y'^2)}} - \frac{\sqrt{1+y'^2}}{\sqrt{2gy}} = \frac{1}{\sqrt{2gy(1+y'^2)}},$$

co po podniesieniu do kwadratu daje

$$y(1+y'^2) = \frac{1}{2gc_1^2} = c.$$

Rozwiązując równanie $y\left(1+\left(\frac{dy}{dx}\right)^2\right) = c$ (o zmiennych rozdzielonych) względem y otrzymamy następujące całki

$$\int \sqrt{\frac{y}{c-y}} dy = \int dx.$$

Zastosujmy podstawienie $y = csin^2\frac{\theta}{2}$.

Zastosujmy podstawienie $y = csin^2\frac{\theta}{2}$. Wówczas $dy = csin\frac{\theta}{2}cos\frac{\theta}{2}d\theta$, a zatem

$$\int dx = \int \sqrt{\frac{csin^2\frac{\theta}{2}}{c - csin^2\frac{\theta}{2}}}c \cdot sin\frac{\theta}{2} \cdot cos\frac{\theta}{2}d\theta = c\int sin^2\frac{\theta}{2}d\theta.$$

Zastosujmy podstawienie $y = csin^2\frac{\theta}{2}$. Wówczas $dy = csin\frac{\theta}{2}cos\frac{\theta}{2}d\theta$, a zatem

$$\int dx = \int \sqrt{\frac{csin^2\frac{\theta}{2}}{c - csin^2\frac{\theta}{2}}} c \cdot sin\frac{\theta}{2} \cdot cos\frac{\theta}{2} d\theta = c \int sin^2\frac{\theta}{2} d\theta.$$

Stosując całkowanie przez części obliczmy

$$\int sin^2 x dx = \int sinx \cdot sinx dx = -sinx \cdot cosx + \int cos^2 x dx =$$

$$-sinx \cdot cosx + \int 1 - sin^2 x dx,$$

Zastosujmy podstawienie $y = csin^2\frac{\theta}{2}$. Wówczas $dy = csin\frac{\theta}{2}cos\frac{\theta}{2}d\theta$, a zatem

$$\int dx = \int \sqrt{\frac{csin^2\frac{\theta}{2}}{c - csin^2\frac{\theta}{2}}} c \cdot sin\frac{\theta}{2} \cdot cos\frac{\theta}{2} d\theta = c \int sin^2\frac{\theta}{2} d\theta.$$

Stosując całkowanie przez części obliczmy

$$\int sin^2 xdx = \int sinx \cdot sinxdx = -sinx \cdot cosx + \int cos^2 xdx =$$

$$-sinx \cdot cosx + \int 1 - sin^2 xdx,$$

stąd

$$\int sin^2 xdx = -\frac{1}{2} sinx \cdot cosx + \frac{x}{2} + d_1.$$

$$\int sin^2 x dx = -\frac{1}{2} sinx \cdot cosx + \frac{x}{2} + d_1,$$

a zatem

$$c \int sin^2 \frac{\theta}{2} d\theta = 2c \int sin^2 x dx = \frac{c}{2}(\theta - sin\theta) + d_2,$$

co daje nam rozwiązanie

$$x = \frac{c}{2}(\theta - sin\theta) + d_3.$$

$$\int sin^2 x dx = -\frac{1}{2} sinx \cdot cosx + \frac{x}{2} + d_1,$$

a zatem

$$c \int sin^2 \frac{\theta}{2} d\theta = 2c \int sin^2 x dx = \frac{c}{2}(\theta - sin\theta) + d_2,$$

co daje nam rozwiązanie

$$x = \frac{c}{2}(\theta - sin\theta) + d_3.$$

Jeśli przyjmiemy $\theta = 0$ dla $x = 0$, to otrzymamy $d_2 = 0$. Ostatecznie korzystając z wzoru $cos^2 x = \frac{cos2x+1}{2}$ dostajemy

$$y = csin^2 \frac{\theta}{2} = c\left(1 - \frac{cos\theta + 1}{2}\right) = \frac{c}{2}\left(1 - cos\theta\right).$$

W efekcie uzyskaliśmy następujące parametryczne równania na współrzędne x i y

$$\begin{cases} x = \frac{c}{2}(\theta - sin\theta) \\ y = \frac{c}{2}\left(1 - cos\theta\right), \end{cases}$$

które są zarazem parametrycznymi równaniami cykloidy, czyli krzywej, jaką zakreśla punkt znajdujący się na obwodzie toczącego się koła ($\frac{c}{2}$ jest promieniem tego koła, natomiast t jest parametrem rzeczywistym odpowiadającym kątowi o jaki obróciło się koło).

W efekcie uzyskaliśmy następujące parametryczne równania na współrzędne x i y

$$\begin{cases} x = \frac{c}{2}(\theta - sin\theta) \\ y = \frac{c}{2}\left(1 - cos\theta\right), \end{cases}$$

które są zarazem parametrycznymi równaniami cykloidy, czyli krzywej, jaką zakreśla punkt znajdujący się na obwodzie toczącego się koła ($\frac{c}{2}$ jest promieniem tego koła, natomiast t jest parametrem rzeczywistym odpowiadającym kątowi o jaki obróciło się koło).
Dowiedliśmy zatem, że jeżeli istnieje krzywa, po której czas staczania się punktu materialnego pod wpływem siły grawitacji jest najmniejszy, to jest to fragment łuku cykloidy. Pokazanie, iż jest to faktycznie minimum, jest dość skomplikowane. Z tego powodu również wiele autorów milczącą pomija tenże dowód, którym my również się nie zajmiemy.

Wyprowadzimy teraz prawo Snelliusa korzystając z zasady Fermata, mówiącej, iż czas jaki zajmuje promieniowi świetlnemu przejście między dwoma ustalonymi punktami, jest równy ekstremum czasu po wszystkich możliwych drogach łączących te punkty.

Wyprowadzimy teraz prawo Snelliusa korzystając z zasady Fermata, mówiącej, iż czas jaki zajmuje promieniowi świetlnemu przejście między dwoma ustalonymi punktami, jest równy ekstremum czasu po wszystkich możliwych drogach łączących te punkty.

W większości przypadków wspomnianym ekstremum z zasady Fermata jest minimum. Załóżmy więc, że mamy jeden z tych przypadków i będziemy szukać minimum.

Wyprowadzimy teraz prawo Snelliusa korzystając z zasady Fermata, mówiącej, iż czas jaki zajmuje promieniowi świetlnemu przejście między dwoma ustalonymi punktami, jest równy ekstremum czasu po wszystkich możliwych drogach łączących te punkty.

W większości przypadków wspomnianym ekstremum z zasady Fermata jest minimum. Załóżmy więc, że mamy jeden z tych przypadków i będziemy szukać minimum.
Analogicznie jak w przypadku zagadnienia brachistochrony będziemy zatem szukać minimum funkcjonału

$$T(y) = \int_a^b t'(x)dx = \int_a^b \frac{\sqrt{1 + y'2}}{v(x, y)} dx,$$

gdzie $v(x, y)$ jest prędkością światła w danym ośrodku o współczynniku załamania światła $n(x, y)$.

Wiemy jednak, że $v(x, y) = \frac{c}{n(x,y)}$, gdzie c jest prędkością światła w próżni. Minimum naszego funckjonału będzie zatem równe minimum funkcjonału

$$\bar{T}(y) = \int_a^b n(x, y)\sqrt{1 + y'2}dx.$$

Wiemy jednak, że $v(x,y) = \frac{c}{n(x,y)}$, gdzie c jest prędkością światła w próżni. Minimum naszego funckjonału będzie zatem równe minimum funkcjonału

$$\bar{T}(y) = \int_a^b n(x,y)\sqrt{1 + y'2}dx.$$

Załóżmy teraz, że n zależy tylko od zmiennej y. Wówczas możemy skorzystać z uproszczonej równości Eulera, czyli dla przypomnienia

$$y' \frac{\partial F}{\partial y'} - F = const.$$

Wiemy jednak, że $v(x,y) = \frac{c}{n(x,y)}$, gdzie c jest prędkością światła w próżni. Minimum naszego funckjonału będzie zatem równe minimum funkcjonału

$$\bar{T}(y) = \int_a^b n(x,y)\sqrt{1+y'2}dx.$$

Załóżmy teraz, że n zależy tylko od zmiennej y. Wówczas możemy skorzystać z uproszczonej równości Eulera, czyli dla przypomnienia

$$y'\frac{\partial F}{\partial y'} - F = const.$$

Po nietrudnych przekształceniach otrzymujemy

$$\frac{n(y)}{\sqrt{1+y'^2}} = const.$$

$$\frac{n(y)}{\sqrt{1+y'^2}} = const.$$

Gdy rozważymy geometryczną interpretację pochodnej, to zauważamy, iż mianownik powyższego ułamka jest równy

$$\sqrt{1+tg^2\theta} = \sqrt{1+\frac{sin^2\theta}{\cos^2\theta}} = \frac{1}{cos\theta},$$

gdzie θ jest kątem pomiędzy styczną do krzywej $y(x)$ a osią OX.

$$\frac{n(y)}{\sqrt{1+y'^2}} = const.$$

Gdy rozważymy geometryczną interpretację pochodnej, to zauważamy, iż mianownik powyższego ułamka jest równy

$$\sqrt{1+tg^2\theta} = \sqrt{1+\frac{sin^2\theta}{\cos^2\theta}} = \frac{1}{cos\theta},$$

gdzie θ jest kątem pomiędzy styczną do krzywej $y(x)$ a osią OX.Możemy zatem naszą równość Eulera zapisać jako

$$n(y)cos\theta = const.$$

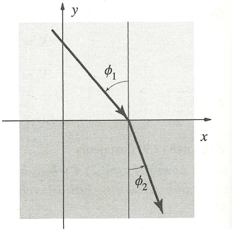
$$n(y) cos\theta = const.$$

$$n(y)cos\theta = const.$$

Pamiętając, że $cos(\frac{\pi}{2} - \phi) = sin\phi$ oraz na mocy powyższej równości dostajemy prawo Snelliusa : jeśli promień światła przechodzi z ośrodka o współczynniku załamania n_1 do ośrodka załamania n_2, to

$$n_1 sin\phi_1 = n_2 sin\phi_2,$$

gdzie ϕ_i, $i = 1, 2$ jest odpowiednim kątem jaki tworzy droga promienia z osią pionową (mowa o prostej prostopadłej do granicy ośrodków).




ZA UWAGĘ
DZIĘKUJĘ
Memy.pl