Wykład 4

Chciałbym dzisiaj wrócić do zagadnienia dotyczącego ogólnie rachunku wariacyjnego, na które jakoś nie było miejsca, bo chciałem jak najszybciej po uzyskaniu równań Eulera przejść do zagadnień dynamiki.

Chodzi o badanie ekstremów warunkowych. Zajmowaliśmy się wcześniej zagadnieniem **rozciągliwej** wiotkiej liny zawieszonej w polu ciężkości, jako przykładem zagadnienia w którym funkcjonał zależy od dwóch różnych funkcji zmiennej niezależnej (a które nie jest zadaniem dynamicznym z czasem jako zmienną niezależną). Po rozwiązaniu tego (dość przydługiego) zadania, zauważyliśmy, że (o ile tylko spoczynkowa długość liny jest większa od odległości punktów zawieszenia) istnieje gładkie przejście do granicy $E \to \infty$, czyli do granicy nici nierozciągliwej. Opis kształtu liny staje się prostszy.

Czujemy intuicyjnie, że zagadnienie kształtu zawieszonej liny **nierozciągliwej**, powinno dać się rozwiązać bez kłopotliwego przechodzenia przez etap liny rozciągliwej. Rzeczywiście jest to możliwe.

Przyjrzyjmy się na początek, jak rozwiązać zagadnienie ekstremum zwykłej funkcji wielu zmiennych $F(x_i)$, jednak nie ekstremum **w ogóle** w całej przestrzeni R^f, a w podzbiorze tej przestrzeni R^f określonej równaniem:

$$G(x_i) = \mathrm{const} \qquad (4.1)$$

Jeśli punkt x^0 jest takim punktem (powiedzmy minimum), to odejście od niego **ale bez schodzenia z warstwicy** $G(x_i) = G(x^0{}_i)$, musi **podwyższyć** wartość funkcji F. Odejście „w bok” może robić cokolwiek. Ekstremum **warunkowe** wymaga by funkcja rosła przy ruchu wzdłuż (uogólnionej) „warstwicy”. Warunek dla różniczek zmiennych niezależnych pozostawania na właściwym podzbiorze sprowadza się do liniowego związku:

$$\sum \frac{\partial G}{\partial x_i} dx_i = 0 \qquad (4.2)$$

Dla różniczek spełniających (4.1) **liniowa część przyrostu** funkcji F, czyli: $\mathrm{d}F = \sum \frac{\partial F}{\partial x_i} dx_i$, **musi być równa zeru:**

$$\sum \frac{\partial F}{\partial x_i} \mathrm{d}x_i = 0 \tag{4.3}$$

Ponieważ teraz różniczki $\mathrm{d}x_i$ **nie są niezależne**, nie można argumentować, że można wybrać takie różniczki, że tylko, na przykład, $\mathrm{d}x_1$ jest różne od zera, a pozostałe zerem, więc znikanie sumy (4.2) sprowadza się do znikania pojedynczego składnika, czyli do znikania $\frac{\partial F}{\partial x_1}$ i tak **po kolei** wszystkich pochodnych cząstkowych, co jest standardowym warunkiem koniecznym ekstremum funkcji wielu zmiennych.

Wszystkie f różniczek nie jest niezależnych, ale $f - 1$ jest! Wszak jest tylko jeden na nie warunek liniowy. Wyliczmy z tego warunku jedną z różniczek (np. pierwszą[1]):

$$\mathrm{d}x_1 = -\sum_{i=2} \frac{\partial G}{\partial x_i} \mathrm{d}x_i \,/\, \frac{\partial G}{\partial x_1} \tag{4.4}$$

Jeśli pierwsza różniczka jest rozumiana jako wyrażenie (4.4), to na różniczki o numerach od 2 do f nie ma **żadnych** ograniczeń.

Wyraźmy liniową część przyrostu funkcji F przez niezależne różniczki:

$$\mathrm{d}F = \frac{\partial F}{\partial x_1}\mathrm{d}x_1 + \sum_{i=2}\frac{\partial F}{\partial x_i}\mathrm{d}x_i = -\frac{\partial F}{\partial x_1}\cdot\sum_{i=2}\frac{\partial G}{\partial x_i}\mathrm{d}x_i \,/\, \frac{\partial G}{\partial x_1} + \sum_{i=2}\frac{\partial F}{\partial x_i}\mathrm{d}x_i =$$

$$= \sum_{i=2}\left(\frac{\partial F}{\partial x_i} - \frac{\partial G}{\partial x_i}\frac{\frac{\partial F}{\partial x_1}}{\frac{\partial G}{\partial x_1}}\right)\mathrm{d}x_i \tag{4.5}$$

[1] Za chwilę będziemy potrzebowali, by $\partial G / \partial x_1$ było różne od zera w punkcie ekstremum. Jeśli **przypadkowo** pochodna ta miałaby znikać, wyróżnimy inną zmienną i wyrazimy jej różniczkę przez pozostałe. **Wszystkie** pochodne nie mogą znikać, w każdym razie tożsamościowo, bo G musiałoby być stałą i nie określało by żadnej podprzestrzeni.

W powyższym wyrażeniu wszystkie $f - 1$ różniczek jest niezależnych, więc warunek ekstremum oznacza, że każdy z nawiasów (**dla *i* = 2,3,...*f***) jest równy zeru:

$$\frac{\partial F}{\partial x_i} - \frac{\dfrac{\partial F}{\partial x_1}}{\dfrac{\partial G}{\partial x_1}} \frac{\partial G}{\partial x_i} = 0 \qquad (4.6)$$

Równań tych jest $f - 1$. Razem z równaniem (4.1) mamy tyle równań ile trzeba dla znalezienia f wartości zmiennych niezależnych.

Równania (4.6) wyglądają i niezbyt elegancko i zasadzają się na nie znikaniu pochodnej cząstkowej po pierwszej zmiennej. Można im nadać znacznie bardziej estetyczną postać, jeśli wprowadzić jeszcze jedną wielkość, *a priori* nieznaną, którą oznaczymy λ i napiszemy dla niej równanie:

$$\lambda = \frac{\dfrac{\partial F}{\partial x_1}}{\dfrac{\partial G}{\partial x_1}} \qquad (4.7)$$

Naturalne jest przepisać do równanie w postaci:

$$\frac{\partial F}{\partial x_1} - \lambda \frac{\partial G}{\partial x_1} = 0 \qquad (4.8)$$

Wraz z definicją (4.7) główne równania 4.6 (dla $i = 2,3,...f$) przyjmują postać:

$$\frac{\partial F}{\partial x_i} - \lambda \frac{\partial G}{\partial x_i} = 0 \qquad (4.9)$$

Równanie (4.8) to nic innego, jak równanie (4.9) dla indeksu $i = 1$ wcześniej wykluczonego. Łącznie z równaniem (4.1) mamy ostatecznie $f + 1$ równań na f niewiadomych, którymi są współrzędne ekstremum i wielkość λ. Równania (4.1) i (4.9) stanowią rozwiązanie problemu ekstremum warunkowego. Równania traktują wszystkie współrzędne na równych prawach.

Wielkość λ, której sens określony jest równaniem (4.7), ale którą wyznaczmy z definicji tej nie korzystając, nazywa się **mnożnikiem Lagrange'a.**

Równanie (4.9) można też uzyskać argumentując **odrobinę** inaczej.

Dla różniczek spełniających wiąz (4.2) (no i dla punktu x_i^0, w którym liczymy pochodne, będącego ekstremum) spełnione są dwa równania niezależnie

$$\sum \frac{\partial G}{\partial x_i} \mathrm{d}x_i = 0, \ \sum \frac{\partial F}{\partial x_i} \mathrm{d}x_i = 0.$$

A zatem (dla różniczek spełniających wiąz, a więc liniowo zależnych), **tożsamościowo ze względu na λ,** spełnione jest równanie:

$$\begin{gathered} \sum \frac{\partial F}{\partial x_i} \mathrm{d}x_i - \lambda \sum \frac{\partial G}{\partial x_i} \mathrm{d}x_i \equiv 0 = \\ \equiv \left(\frac{\partial F}{\partial x_1} - \lambda \frac{\partial G}{\partial x_1} \right) \mathrm{d}x_1 + \left(\frac{\partial F}{\partial x_2} - \lambda \frac{\partial G}{\partial x_2} \right) \mathrm{d}x_2 + \cdots \left(\frac{\partial F}{\partial x_f} - \lambda \frac{\partial G}{\partial x_f} \right) \mathrm{d}x_f \end{gathered} \tag{4.10}$$

Różniczki nie są niezależne, ale $f - 1$ spośród nich jest! Możemy więc w szczególności położyć wszystkie różniczki o wskaźniku od 3 do f równe 0, $\mathrm{d}x_2$ **różne** od zera. Mamy:

$$\left(\frac{\partial F}{\partial x_1} - \lambda \frac{\partial G}{\partial x_1} \right) \mathrm{d}x_1 + \left(\frac{\partial F}{\partial x_2} - \lambda \frac{\partial G}{\partial x_2} \right) \mathrm{d}x_2 \equiv 0 \tag{4.11}$$

Różniczkę $\mathrm{d}x_1$ można by wyliczać, ale nam się nie chce! Ona jest jakaś, akurat tak wyrażona przez $\mathrm{d}x_2$, że dwa człony, jakie pozostają **tożsamościowo,** (tj. niezależnie od wartości λ) skracają się. O samych (teraz 2) nawiasach nic mądrego nie mogę powiedzieć.

Skoro λ może być **czymkolwiek**, to można w szczególności wybrać je tak, by znikał pierwszy nawias. Znika na mocy wyboru mnożnika Lagrange'a, który już przestaje być byle czym! Równanie (4.11) sprowadza się **teraz** (już dla określonego λ) do równania:

$$\left(\frac{\partial F}{\partial x_2} - \lambda \frac{\partial G}{\partial x_2} \right) = \frac{\partial (F - \lambda G)}{\partial x_2} = 0 \tag{4.12}$$

Zamiast wskaźnika 2 możemy wziąć, z kolei wskaźnik 3. Pierwszy nawias, dla ustalonego już λ, ciągle jest zerem, a w miejsce równania (4.12) dostaniemy po kolei wszystkie równania (4.9) dla ekstremum warunkowego.

To drugie rozumowanie łatwo jest teraz uogólnić na dowolną liczbę (byle mniejszą od f), warunków dodatkowych. Przykładowo, dla dwóch:

$$\sum \frac{\partial G_1}{\partial x_i} dx_i = 0, \quad \sum \frac{\partial G_2}{\partial x_i} dx_i = 0 \tag{4.13}$$

tylko $f - 2$ różniczek można wybrać dowolnie. Pozostałe 2 można wyznaczyć rozwiązując układ powyższych dwóch równań. Póki różniczki spełniają oba więzy, w punkcie ekstremum, zachodzi tożsamościowo dla dowolnych wartości pary λ_1, λ_2

$$\begin{aligned} &\sum \frac{\partial F}{\partial x_i} \mathrm{d}x_i - \lambda_1 \sum \frac{\partial G_1}{\partial x_i} \mathrm{d}x_i - \lambda_2 \sum \frac{\partial G_2}{\partial x_i} \mathrm{d}x_i \equiv 0 = \\ &\equiv \left(\frac{\partial F}{\partial x_1} - \lambda_1 \frac{\partial G_1}{\partial x_1} - \lambda_2 \frac{\partial G_2}{\partial x_1} \right) \mathrm{d}x_1 + \left(\frac{\partial F}{\partial x_2} - \lambda_1 \frac{\partial G_1}{\partial x_2} - \lambda_2 \frac{\partial G_2}{\partial x_2} \right) \mathrm{d}x_2 + \qquad 4.14) \\ &+ \left(\frac{\partial F}{\partial x_2} - \lambda_1 \frac{\partial G_1}{\partial x_3} - \lambda_2 \frac{\partial G_2}{\partial x_3} \right) \mathrm{d}x_3 + \cdots + \left(\frac{\partial F}{\partial x_f} - \lambda_1 \frac{\partial G_1}{\partial x_f} - \lambda_2 \frac{\partial G_2}{\partial x_f} \right) \mathrm{d}x_f \end{aligned}$$

Jeśli zamiast cieszyć się, że oba mnożniki Lagrange'a mogą być dowolne, wybierzemy konkretne 2 **zerujące** pierwsze dwa nawiasy (w punkcie ekstremum), to pozostała suma (dla tych konkretnych) musi znikać już dla **dowolnych** $f - 2$ pozostałych różniczek, bo są one przecież niezależne. Ich ustalenie (dowolne) fiksuje pozostałe 2 różniczki, ale nawiasy przy nich są już zerami.

Mamy więc układ równań f:

$$\frac{\partial (F - \lambda_1 G_1 - \lambda_2 G_2)}{\partial x_i} = 0, \tag{4.15}$$

uzupełniony rzecz jasna dwoma dalszymi: $G_1(x) = C_1, G_2(x) = C_{2w}$

Ostatecznie dochodzimy do prostego algorytmu. Chcąc znaleźć ekstremum warunkowe F przy warunkach

$$G_1(x) = C_1, G_2(x) = C_2 \cdots G_w(x) = C_w, \tag{4.16}$$

tworzymy funkcję:

$$F(x)-\lambda_1 G_1(x)-\lambda_2 G_2(x)\cdots-\lambda_w G_w(x)\,, \tag{4.16}$$

w której mnożniki Lagrange'a traktujemy jako stałe parametry i dla której wypisujemy zwykłe warunki na ekstremum **bezwarunkowe** (to znaczy przyrównujemy do zera wszystkie pochodne cząstkowe względem f zmiennych niezależnych.

Algorytm powyższy nie zależy od liczby zmiennych. Dla funkcjonałów, które można sobie wyobrażać jako graniczny przypadek funkcji **nieskończenie wielu** zmiennych, działa to dokładnie tak samo!

Teraz już możemy wziąć się za naszą krzywą łańcuchową. Warunkiem równowagi jest zwykłe minimum energii potencjalnej grawitacji:

$$\cancel{\rho g S}\int y(x)\sqrt{1+y'^2}\,\mathrm{d}x=\min \tag{4.17}$$

przy warunku ubocznym:

$$\int\sqrt{1+y'^2}\,\mathrm{d}x=l \tag{4.18}$$

Mamy jeden warunek, zatem jeden mnożnik Lagrange'a. Należy wypisać warunek minimum dla odpowiedniej różnicy:

$$\delta\int(y(x)-\lambda)\sqrt{1+y'^2}\,\mathrm{d}x=0 \tag{4.19}$$

Tworzymy całkę pierwszą:

$$y'\frac{\partial(y-\lambda)\sqrt{1+y'^2}}{\partial y'}-(y-\lambda)\sqrt{1+y'^2}=C \tag{4.20}$$

$$(y-\lambda)\left(\frac{y'^2}{\sqrt{1+y'^2}}-\sqrt{1+y'^2}\right)=C \tag{4.21}$$

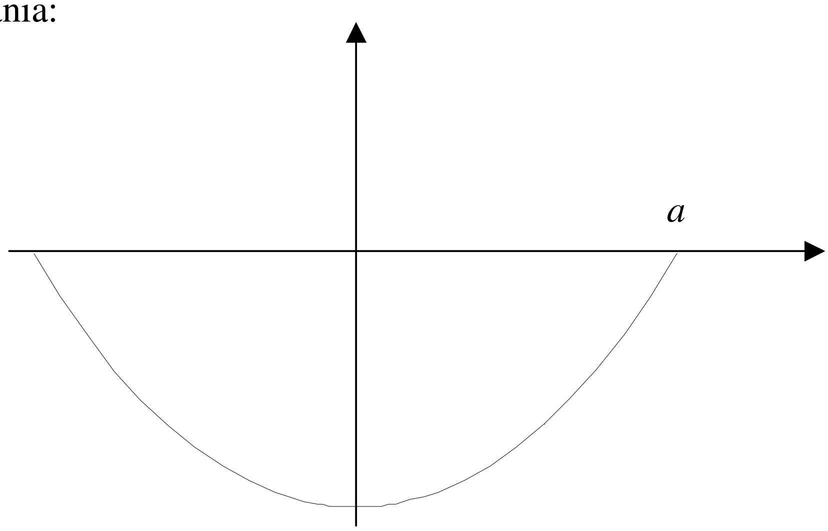
$$(y-\lambda)\left(\frac{-1}{\sqrt{1+y'^2}}\right)=C \tag{4.22}$$

$$y'^2 = \left(\frac{y-\lambda}{C}\right)^2 - 1 \tag{4.23}$$

$$x = \int \frac{\mathrm{d}y}{\sqrt{\left(\frac{y-\lambda}{C}\right)^2 - 1}} = C\,\mathrm{Arch}\,\frac{y-\lambda}{C} + D \tag{4.24}$$

$$y = \lambda + C\cosh\frac{x}{C} \tag{4.25}$$

Jeśli początek układu wybierzemy tak, by wierzchołek odpowiadał wartości $x = 0$, to $D = 0$. Jeśli ponadto, współrzędną y mierzymy od poziomu zamocowania:



czyli żądamy by: $y(a) = 0$, to $0 = \lambda + C\cosh\frac{a}{C}$ i mamy:

$$y = C\left(\cosh\frac{x}{C} - \cosh\frac{a}{C}\right) \tag{4.26}$$

Ostatnią stałą, jaka pozostała wyznaczymy z warunku długości nici ($2l$):

$$\begin{aligned} y' &= \sinh\frac{x}{C} \\ l &= \int_0^a \sqrt{1+\sinh^2\frac{x}{C}}\,\mathrm{d}x = C\sinh\frac{a}{C} \end{aligned} \tag{4.27}$$

Przy danej długości ($2l$) i rozstawieniu ($2a$), równanie (4.27) trzeba rozwiązać numerycznie.

Dysponując dzisiejszymi możliwościami, rozwiązywanie równań przestępnych, to żaden problem. Czasami jednak potrzebny nam jest wzór ogólny, zachowanie asymptotyczne, czy też akurat pod ręką nie mamy komputera.

Równanie (4.27) daje dobrą okazję do przedstawienia pewnej ogólnej metody rozwiązywania równań przestępnych metodą kolejnych przybliżeń. Przed przejściem do rozważań numerycznych warto wprowadzić wielkości bezwymiarowe. Dzieląc równanie (4.27) przez *a* mamy:

$$l / a = \frac{C}{a} \sinh \frac{a}{C} \tag{4.28}$$

Oznaczmy *l/a* symbolem *d* (długość w jednostkach wartości rozpiętości punktów zawieszenia), a bezwymiarową wielkość niewiadomą symbolem *x:*

$$a/C = x \tag{4.29}$$

Do rozwiązania mamy:

$$\sinh x = x \cdot d \tag{4.30}$$

Funkcja sinh*x*/*x* jest monotonicznie rosnąca od 1 do nieskończoności, zatem im większe *d*, tym większy pierwiastek *x.*

Gdy długość (względem rozpiętości) niewiele przewyższa wartość 1, rozwiązanie *x* jest małe i wystarczy rozwinąć sinus hiperboliczny zachowując dwa pierwsze wyrazy:

$$\begin{aligned} & x + \frac{1}{6} x^3 \cong x \cdot d \\ & x = \sqrt{6(d-1)} \end{aligned} \tag{4.40}$$

Znając *x,* możemy odpowiedzieć, w szczególności, o ile obniży się środek liny:

$$-y(0) = C(\cosh(a/C) - 1) \cong \frac{1}{2} a^2 / C = \frac{1}{2} a x = a \sqrt{\frac{3}{2}\left(\frac{l}{a} - 1\right)} \tag{4.41}$$

Gdy długość liny staje się „nieduże” parę razy większa od odległości punktów zamocowania, rozwinięcie do dwóch wyrazów staje się za mało dokładne i formuły analitycznej nie ma pod ręką.

Ciekawa sytuacja powstaje wtedy, gdy lina jest dużo dłuższa od odstępu punktów zawieszenia. Pojawia się charakterystyczna metoda postępowania z równaniami postaci $x = f(x)$. W naszym przypadku piszemy:

$$x = \mathrm{Arsh}(x \cdot d) \qquad (4.42)$$

Dlaczego tak? To za chwilę stanie się jasne. Rzecz w tym, że dla dalekich argumentów (a x rośnie ze wzrostem d), pochodna funkcji Arsh(x) jest **mała.** W znacznym obszarze wokół właściwej wartości x prawa strona równania pozostaje niewiele różna (bo wolno się zmienia) od wartości właściwej. Inaczej mówiąc, ciąg liczb:

$$x_{n+1} = \mathrm{Arsh}(x_n \cdot d) \qquad (4.43)$$

stanowi zbieżny ciąg przybliżeń rozwiązania równania (4.42)

Dla ilustracji wezmę wartość $d = 10$.

```
NestList[ArcSinh[10 #] &, x, 4]
 {x, ArcSinh[10 x], ArcSinh[10 ArcSinh[10 x]],
  ArcSinh[10 ArcSinh[10 ArcSinh[10 x]]],
  ArcSinh[10 ArcSinh[10 ArcSinh[10 ArcSinh[10 x]]]]}
```

Widzimy jak działa funkcja NestList. Tworzy dokładnie to co chcemy, czyli ciąg (4.43).

A teraz utwórzmy taki ciąg (nieco dłuższy, ale już liczbowy) zaczynając od ...??? no byle czego. Nawet od jakiejś głupoty, powiedzmy 20:

```
NestList[ArcSinh[10 #] &, 20., 10]
 {20., 5.99147, 4.78614, 4.56157, 4.51352,
  4.50293, 4.50058, 4.50006, 4.49995, 4.49992, 4.49992}
```

Zupełnie błędny „guess” 20 (zamiast 4.5) już po pierwszej iteracji jest sprowadzony do „rozsądnych” granic, tj. 6, by w następnym kroku dać 4.79, po czym już 4.56 itd. Dziewiąty wyraz ciągu (4.49992) podstawiony do prawej strony równania daje **to samo**. Równanie jest więc rozwiązane.

Można łatwo się przekonać, że błędy takiego ciągu przybliżeń spadają jak postęp geometryczny o ilorazie równym pochodnej funkcji f występującej w równaniu. Pozornie równe dobre do iteracji równanie

$$x = \frac{1}{d}\sinh x \tag{4.44}$$

prowadzi na manowce, nawet gdy zacznę od doskonałego przybliżenia!!

```
NestList[0.1 Sinh[#] &, 4.4999, 10]
```

```
{4.4999, 4.49985, 4.49963, 4.49864, 4.49417,
 4.47415, 4.38544, 4.01304, 2.76483, 0.790665, 0.087566}
```

Gdy zacznę byle jak, program się gniewa:

a
```
NestList[0.1 Sinh[#] &, 10, 10]
```

```
General::ovfl : Overflow occurred in computation.
```

$\{10,\ 1101.32,\ 9.94481123887\times10^{476},\ \mathrm{Overflow}[],$
Indeterminate, Indeterminate, Indeterminate,
Indeterminate, Indeterminate, Indeterminate, Indeterminate}

Pierwiastek równania

$x = f(x)$

Jest też pierwiastkiem równania z funkcją odwrotną:

$x = f^{-1}(x)$

Pochodna funkcji odwrotnej jest równa odwrotności pochodnej. Gdy funkcja f jest szybko rosnąca (jak sinus hiperboliczny), czyli ma dużą pochodną w okolicach rozwiązania, to funkcja odwrotna ma pochodną **małą.** O czym przekonaliśmy się na przykładzie.