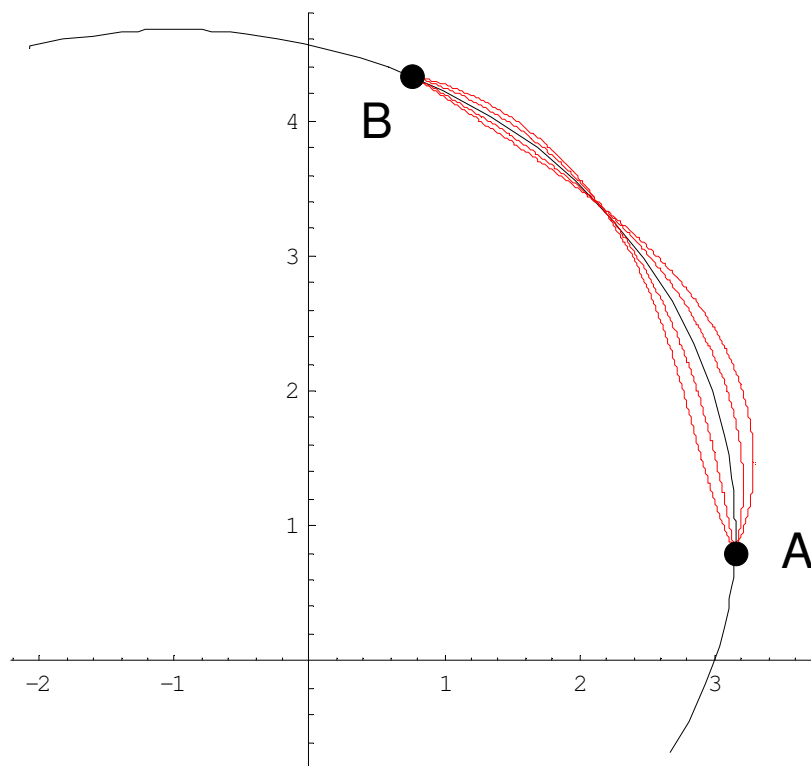


Wykład 1

(3 godz.)

Zacznę od przypomnienia pewnego rysunku i pewnego rachunku z wykładu zeszłorocznego.



Na czarno zaznaczony jest rzut toru punktu materialnego zmuszonego do poruszania się po powierzchni obrotowej (bez działania innych sił niż stwarzane przez powierzchnię siły **reakcji**). Mimo, iż siła reakcji nie jest znana **z góry**, sam fakt, że jest ona prostopadła do powierzchni, a jej rzut na płaszczyznę rysunku jest centralny, pozwalał napisać prawo zachowania energii i momentu pędu, co wystarczyło nam do wyznaczenia równania tego toru.

Dla linii leżącej na zadanej powierzchni i łączącej dwa zadane z góry punkty możemy postawić też, czysto geometryczne, zadanie znalezienia takiej linii, której długość jest **ekstremalna**, w tym wypadku **najkrótsza** z możliwych.

Pamiętacie Państwo, że fizyczny tor jest równocześnie ekstremalą. Uczciwie to policzyłem. Pominę sposób wyprowadzenia równania toru z praw fizyki (w tym wypadku praw zachowania), przypomnę wyprowadzenie równania wynikającego z warunku ekstremalności.

Na rysunku są jakieś dwa punkty na torze cząstki poruszającej się po powierzchni bryły obrotowej, jest fragment stosownej geodezyjnej i kilka **alternatywnych** dróg, razem pięć, wiodących od A do B. Dokładniej mówiąc, owe pięć dróg odpowiada zależnościom $\varphi(r) = \varphi_{\text{geo}}(r) + \varepsilon \delta(r)$, dla $\varepsilon = 1, 0.5, 0, -0.5, -1$, gdzie $\varphi_{\text{geo}}(r)$ **jest fragmentem geodezyjnej** (dla jakiejś metryki, której nie specyfikuję, bo rozważania są ogólne). Z kolei $\delta(r)$ jest

jakaś, jedną z nieskończenie wielu możliwych, modyfikacją, (naukowo zwaną wariacją), naszej geodezyjnej. **Oczywistą** własnością $\delta(r)$, jeśli mamy porównywać długość odcinka geodezyjnej z długością alternatywnych dróg **od A do B**, jest jej **znikanie** w punktach A i B

$$\delta(r_A) = \delta(r_B) = 0$$

Nie jest trudno wyobrazić sobie alternatywne drogi **dla wszystkich** wartości ε pomiędzy -1 , a $+1$. Oczywiście długość drogi staje się w tej sytuacji funkcją ε .

Cokolwiek wybralibyśmy na funkcję $\delta(r)$, wartość $\varepsilon = 0$, będzie odpowiadała akurat geodezyjnej (w naszym rozumieniu tego określenia, jako drogi wybieranej przez punkt materialny, nie doznający żadnej siły w płaszczyźnie stycznej)

$$l_{AB}(\varepsilon) = \int_{r_A}^{r_B} \sqrt{r^2 d\varphi^2 + g_{rr}(r) dr^2} = \int_{r_A}^{r_B} \sqrt{r^2 (\varphi'_{\text{geo}}(r) + \varepsilon \delta'(r))^2 + g_{rr}(r)} dr \quad (1.1)$$

Dla jakiej wartości ε droga jest najkrótsza?

To nie jest bardzo trudne pytanie! Warunkiem koniecznym jest, by **pochodna** po ε była zerem! Domyślamy się, że tak właśnie będzie dla $\varepsilon=0$. Jak niebawem zrozumiemy, fizyczną istotą geodezyjnej, jako toru dla cząstek (albo światła, wszak widzimy że cząstka ultrarelatywistyczna, porusza się na naszej powierzchni po **takim samym** torze jak i cząstka powolna), jest nie tyle jej „**najkrótszość**”, ale właśnie to, by zmiany długości, przy małych wariacjach, były proporcjonalne do wyższych potęg ε niż 1. A to właśnie pochodna mnoży pierwszą potęgę przyrostu funkcji:

$$f(x + \varepsilon) - f(x) = \varepsilon f'(x) + O(\varepsilon^2)$$

W większości prostych sytuacji, dla geodezyjnej (czyli dla $\varepsilon=0$) mamy rzeczywiście minimum długości drogi, ale to akurat nie jest ważne.

Wypada policzyć (dla wartości $\varepsilon = 0$) tę pochodną:

$$\left. \frac{dl_{AB}(\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0} = \int_{r_A}^{r_B} \frac{r^2 \varphi'_{\text{geo}}(r)}{\sqrt{r^2 (\varphi'_{\text{geo}}(r))^2 + g_{rr}(r)}} \delta'(r) dr = \int_{r_A}^{r_B} \frac{r^2 \varphi'_{\text{geo}}(r)}{\sqrt{r^2 (\varphi'_{\text{geo}}(r))^2 + g_{rr}(r)}} d\delta(r) \quad (1.2)$$

Iloczyn pochodnej wariacji $\delta'(r)$ i przyrostu dr zastąpiłem przyrostem $\delta'(r)dr = d\delta(r)$.

Sztuczka z uzupełnianiem adb o bda do pełnego przyrostu $d(ab)$, jest niezwykle uży-

teczna!. Przyrosty pełnego iloczynu $d\left(\frac{r^2 \varphi'_{\text{geo}}(r)}{\sqrt{r^2 (\varphi'_{\text{geo}}(r))^2 + g_{rr}(r)}} \delta(r)\right)$ zsumują się nieuchronnie

do **totalnego** przyrostu, czyli różnicy wartości tego **iloczynu** w punktach B i A. **Ale tam jeden z czynników** $\delta(r_A) = 0 = \delta(r_B) = 0$!, więc i cały iloczyn jest tam **równy zeru**.

Ponieważ pod całką dodaliśmy bda , co z istniejącym wyrażeniem dało ostatecznie, tj. po wycalkowaniu, 0, musimy jeszcze tylko uwzględnić człon $-bda$, (bo „uzupełnianie” to przecież **dodanie i odjęcie** owego bda). Zatem:

$$\left. \frac{dl_{AB}(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = - \int_{r_A}^{r_B} \delta(r) d \frac{r^2 \phi'_{\text{geo}}(r)}{\sqrt{r^2 (\phi'_{\text{geo}}(r))^2 + g_{rr}(r)}} = - \int_{r_A}^{r_B} \delta(r) \frac{d}{dr} \frac{r^2 \phi'_{\text{geo}}(r)}{\sqrt{r^2 (\phi'_{\text{geo}}(r))^2 + g_{rr}(r)}} dr \quad (1.3)$$

Ale popatrzymy¹ na równanie **uzyskane z mechaniki**. Toż mówi ono dokładnie to, iż wyrażenie do różniczkowania:

$$\frac{r^2 \phi'_{\text{geo}}(r)}{\sqrt{r^2 (\phi'_{\text{geo}}(r))^2 + g_{rr}(r)}},$$

jest **stałą**. A więc „nasza” definicja geodezyjnej, jako toru fizycznej cząstki, **najprostszej linii**, pozwala udowodnić, że jest ona nie tylko minimalna pod względem krzywizny, ale i minimalna (czy ogólniej **stacjonarna**) pod względem **długości**. I na odwrót². Jeśli ktoś definiuje geodezyjną jako linię „najkrótszą”, to powyższe rozumowanie dowodzi, że punkty materialne i światło (zmuszone pozostawać na powierzchniach z dowolną metryką) **właśnie będą poruszać się** po takich geodezyjnych.

Zapamiętajmy ten rachunek. Gdy w całce $\int_A^B F(y'(x), x) dx$ nie występuje samo $y(x)$, a jedynie $y'(x)$, to zależność $y(x)$ minimalizująca tę całkę spełniać **musi** równanie $\frac{\partial F(y', x)}{\partial y'} = \text{constans}$

Jest to rewelacyjnie wygodny i skuteczny sposób dochodzenia do potrzebnych równań.

Bezpośredni związek fizycznego toru cząstki na powierzchni więzów z geodezyjną zdefiniowaną jako linia ekstremalna jest zaledwie **czubkiem góry lodowej**. Fizyka, a w szczególności mechanika, aż roi się od tego rodzaju związków. Wygodnie jest studiować charakterystyczne zagadnienia mechaniki teoretycznej znając od początku sformułowanie jej ogólnych równań ruchu, jako równań pewnej zasady **wariacyjnej**.

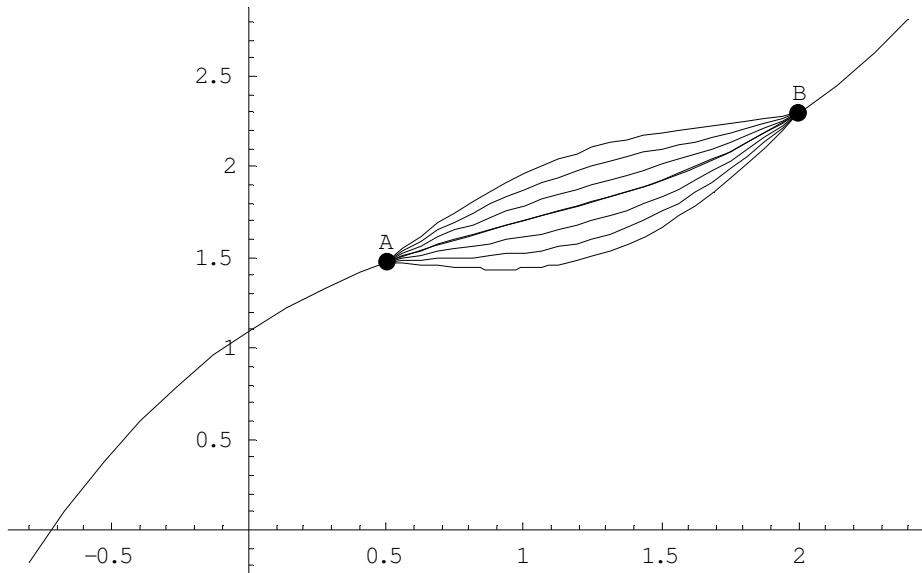
Zacniemy od zbadania aspektu czysto matematycznego zagadnienia wariacyjnego **ogólniejszego** od tego wyniku zawartego w ramach. Pozostając chwilowo przy jednej tylko funkcji zależnej (i jednej zmiennej niezależnej), dopuścimy zależność funkcji podcałkowej nie tylko od x i $y'(x)$, ale także od $y(x)$.

¹ Powinniśmy „popatrzeć” do wykładu 15 z Podstaw Fizyki.

² Jeśli żądamy, by krzywa opisana funkcją $\phi(r)$ była ekstremalą, (czyli albo minimalna, albo maksymalna, albo chociaż stacjonarna) żądać musimy znikania pochodnej ostatniej całki, **dla każdej** funkcji $\delta(r)$, a to wymusza znikanie wzdłuż całego toru pochodnej występującego tam wyrażenia, co oznacza jego stałość.

$$\int_A^B F(y, y'(x), x) dx = \text{ekstremum} \quad (1.4)$$

Zmienne x i y kojarzą się nam raczej z współrzędnymi kartezjańskimi zastąpmy więc pierwotny rysunek takim, jaki najczęściej spotykany jest w standardowych podręcznikach, czy to mechaniki, czy rachunku wariacyjnego:



W przedziale od A do B rozważamy, obok ustalonej funkcji $\bar{y}(x)$ także *continuum* funkcji **porównawczych** $y(x) = \bar{y}(x) + \varepsilon \delta(x)$, gdzie $\delta(A) = \delta(B) = 0$.

Chcemy znaleźć warunek na funkcję $\bar{y}(x)$ (ekstremalę danego funkcjonału), czyli warunek na $\bar{y}(x)$ taki, by dla takiej akurat funkcji, funkcjonał:

$$\int_A^B F(y, y'(x), x) dx = \int_A^B F(\bar{y}(x) + \varepsilon \delta(x), \bar{y}'(x) + \varepsilon \delta'(x), x) dx \quad (1.5)$$

miał **znikającą** pochodną po ε , dla $\varepsilon=0$, dla **każdego** $\delta(x)$.

Zróżniczkujemy więc naszą całkę po ε , połóżmy $\varepsilon=0$ i przyrównajmy wynik do zera. Dostajemy:

$$0 = \int_A^B (\delta(x) F_{,\bar{y}}(\bar{y}(x), \bar{y}'(x), x) + \delta'(x) F_{,\bar{y}'}(\bar{y}(x), \bar{y}'(x), x)) dx \quad (1.6)$$

W dalszych przekształceniach rozróżnienie między ustalonym przebiegiem $\bar{y}(x)$, a zależnym od parametru ε przebiegiem **porównawczym** $y(x)$ nie będzie już potrzebne. Pomińmy niewygodną w drukowaniu kresczkę nad y .

$$0 = \int_A^B (\delta(x) F_{,y}(y(x), y'(x), x) + \delta'(x) F_{,y'}(y(x), y'(x), x)) dx \quad (1.7)$$

W drugim członie całkujemy przez części – człon brzegowy znika, bo znika **wariacja** na brzegach przedziału: $\delta(A) = \delta(B) = 0$. Dostajemy:

$$\begin{aligned} 0 &= \int_A^B \left(\delta(x) F_{,y}(y(x), y'(x), x) - \delta(x) \frac{d}{dx} F_{,y'}(y(x), y'(x), x) \right) dx = \\ &= \int_A^B \left(F_{,y}(y(x), y'(x), x) - \frac{d}{dx} F_{,y'}(y(x), y'(x), x) \right) \delta(x) dx \end{aligned} \quad (1.8)$$

Po to, by powyższa pochodna była zerem dla **każdej** pomyślanej funkcji $\delta(x)$ (znikającej oczywiście na brzegu), znikać musi (tożsamościowo dla wszystkich x) wyrażenie w nawiasie:

$$F_{,y}(y(x), y'(x), x) - \frac{d}{dx} F_{,y'}(y(x), y'(x), x) = 0 \quad (1.9)$$

Zapisujemy to zazwyczaj tak:

$$\frac{d}{dx} \frac{\partial}{\partial y'} F(y(x), y'(x), x) = \frac{\partial}{\partial y} F(y(x), y'(x), x) \quad (1.10)$$

Jest to słynne **równanie Eulera**.

Gdy funkcja podcałkowa rozpatrywanego funkcjonału F nie zależy *explicite* od zmiennej y (a jedynie od jej pochodnej), prawa strona powyższego równania znika, co implikuje stałość wzdłuż ekstremali wielkości $\frac{\partial}{\partial y'} F(y(x), y'(x), x) = \text{const}$. To znamy z poprzedniego wykładu.

W przypadku ogólnym, równanie Eulera jest równaniem różniczkowym **drugiego rzędu**.

W zasadzie Fermata, gdy współczynnik załamania jest funkcją tylko wysokości z , jeśli kształt promienia opiszemy jako funkcję $x(z)$, mamy:

$$\min = \int n(z) \sqrt{(dx)^2 + (dz)^2} = \int n(z) \sqrt{(x'(z))^2 + 1} dz \quad (1.11)$$

co jest tym najprostszym problemem rozpatrywanym przez nas już w zeszłym roku (Rozwiązaniem jest $n(z)x'(z)/\sqrt{(x'(z))^2 + 1} = \text{const}$, co jest r. r. 1-go rzędu).

Jeśli jednak współczynnik załamania zależy zarówno od z jak i od x , ani opis kształtu jako funkcji $z(x)$, ani jako $x(z)$, nie da nam naszego "łatwego" przypadku. Musimy odwołać się do ogólniejszego równania 2-go rzędu:

$$\frac{d}{dz} \frac{n(x(z), z) \cdot x'(z)}{\sqrt{(x'(z))^2 + 1}} = \frac{\partial n(x, z)}{\partial x} \quad (1.12)$$

Wbrew pozorom jest to równanie różniczkowe o pochodnych **zwyczajnych**. Występująca pochodna cząstkowa oznacza różniczkowanie funkcji **znanej**. Po wykonaniu tego różniczkowania, prawa strona jest znaną funkcją zmiennych x i z . Po lewej stronie pojawią się człony zawierające **drugą** pochodną szukanej funkcji $x(z)$. Bez konkretnego kształtu zależności $n(x, z)$, o równaniu tym nic ciekawego już nie można powiedzieć. W razie potrzeby zawsze można rozwiązać je numerycznie.

Interesujący wątek dostaniemy, gdy mając „prosty” współczynnik załamania zależny tylko od z , uprzemy się by kształt promienia opisywać jako funkcję $z(x)$. Dostajemy:

$$\min = \int n(z) \sqrt{(dx)^2 + (dz)^2} = \int n(z) \sqrt{z'^2 + 1} dx \quad (1.13)$$

Funkcja podcałkowa F zależy teraz i od pochodnej (z') i od zmiennej zależnej z . Zastosować wyniku uproszczonego nie możemy, a przecież wiemy, że w problemie tym istnieje, prosto wyrażona, wielkość **stała** wzdłuż promienia.

Zbadajmy wielkość następującą:

$$y'F_{,y'} - F \quad (1.14)$$

i obliczmy jej pochodną **zupełną** po zmiennej niezależnej x .

$$\frac{d}{dx} (y'F_{,y'} - F) = y''F_{,y'} + y' \frac{d}{dx} F_{,y'} - F_{,x} - y''F_{,y'} - y'F_{,y} \quad (1.15)$$

Widzimy, że człon pierwszy redukuje się z przedostatnim. Człon drugi, **na mocy równania Eulera** redukuje się z członem ostatnim i zostaje:

$$\frac{d}{dx} (y'F_{,y'} - F) = -F_{,x} \quad (1.16)$$

Jeśli funkcja F **zależy explicite** od zmiennej niezależnej, mamy całkę pierwszą:

$$F_{,x} = 0 \Rightarrow y'F_{,y'} - F = \text{const} \quad (1.17)$$

W przypadku jednowymiarowym (jedna zmienna zależna) powyższe równanie jest tylko przepisany prawem zachowania starego typu, jeśli zamienić rolami zmienną zależną i niezależną. W przypadku, gdy zmienną niezależną będzie czas, powyższa całka okaże się **całką energii**.

Podsumowując:

Gdy funkcja podcałkowa problemu wariacyjnego zależy od wszystkich trzech argumentów w sposób istotny: $F = F(y, y', x)$, warunek konieczny na ekstremalę jest **równaniem różniczkowym drugiego rzędu** $\frac{d}{dx} \frac{\partial}{\partial y'} F(y(x), y'(x), x) = \frac{\partial}{\partial y} F(y(x), y'(x), x)$

Gdy funkcja podcałkowa **nie zależy** od zmiennej zależnej (a jedynie od jej pochodnej $F = F(y', x)$), równanie Eulera jest równoważne równaniu pierwszego rzędu $F_{,y'} = \text{const}$.

Gdy funkcja podcałkowa nie zależy explicite od zmiennej zależnej, $F = F(y, y')$, równanie Eulera jest równoważne równaniu pierwszego rzędu: $y' F_{,y'} - F = \text{const}$.

Kolejne, dość oczywiste, uogólnienie problemu wariacyjnego polega na zastąpieniu **pojedynczej** zmiennej zależnej y (traktowanej jako funkcja zmiennej niezależnej, którą **od tego miejsca** zaczniemy oznaczać literą t), **kolekcją** takich funkcji $y_i(t)$, gdzie $i=1,2,...,f$. Pochodne po **tej** zmiennej, oznaczać będziemy kropką nad zmienną zależną.

Chodzi więc o funkcjonał:

$$\int F(y_i, \dot{y}_i, t) dt \quad (1.18)$$

Z funkcjonałem takim zetknęliśmy się rok temu, obliczając w szczególnej teorii względności czas własny cząstki opisanej równaniami ruchu:

$$x(t), y(t), z(t).$$

Wynik na ten czas, to

$$\int_{t_A}^{t_B} \sqrt{1 - (\dot{x}^2(t) + \dot{y}^2(t) + \dot{z}^2(t)) / c^2} dt \quad (1.19)$$

Jest zupełnie oczywiste, że jeden warunek znikania wariacji (liniowej części przyrostu) funkcjonału zamieni się teraz na f warunków dla każdej zmiennej zależnej z osobna. No bo jeśli dana trajektoria w $f+1$ wymiarowej przestrzeni jest ekstremalą, to **nie ruszając** $f-1$ zależności y_i od t , a jedynie jedną, też mamy ekstremum. Możemy posłużyć się równaniem Eu-

lera tak jakby $f-1$ zmiennych było zafiksowanych, co sprowadza problem do problemu jednowymiarowego. Ostatecznie mamy:

$$\frac{d}{dt} \frac{\partial}{\partial \dot{y}_i} F(y(x), \dot{y}(x), x) = \frac{\partial}{\partial y_i} F(y(x), \dot{y}(x), x) \text{ dla } i = 1, 2, 3, \dots, f \quad (1.20)$$

Jest to układ f równań różniczkowych drugiego rzędu.

Jeśli postać funkcjonału **nie jest najogólniejsza**, w szczególności, jeśli funkcja F nie zależy od jakiejś zmiennej zależnej (lub od pewnej liczby takich zmiennych), lub jeśli nie zależy od zmiennej niezależnej, możemy podać od razu pewne wielkości zależne tylko od **pierwszych pochodnych** (czyli prędkości, jeśli zmienna t będzie czasem), które są **stałe** wzdłuż rozwiązania (czyli stałe w czasie).

Jeśli $\frac{\partial}{\partial y_i} F(y(t), \dot{y}(t), t) = 0$ dla jakiegoś i , to dla tego samego i : $\frac{\partial}{\partial \dot{y}_i} F(y(t), \dot{y}(t), t) = \text{const}$

Jeśli $\frac{\partial}{\partial t} F(y(t), \dot{y}(t), t) = 0$ (czyli, jeśli $F = F(y, \dot{y})$) to: $\sum \dot{y}_i \frac{\partial F}{\partial \dot{y}_i} - F = \text{const}$

Dowód stałości ostatniego wyrażenia jest niemal identyczny jak w przypadku z jedną zmienną. Pochodne po czasie obu składników to teraz są sumy po i , a poszczególne człony znoszą się kolejno, bo równania Eulera obowiązują dla każdego i z osobna.

Przekonamy się, że znane nam już (a także inne jeszcze nieznanne) całki pędu i momentu pędu są typu pierwszego. Całka energii jest całką drugiego typu.

Dotychczas to była abstrakcyjna matematyka. Jak „zaprząć” rachunek wariacyjny do mechaniki?

Można wyobrazić sobie dwa podejścia.

Przyjmując za punkt wyjścia równania Newtona dla pewnej liczby N punktów materialnych traktowane jako podpatrzone prawa przyrody, a dokładniej biorąc ich szczególny przypadek z siłami potencjalnymi³, i zapisując je w postaci:

$$m^A \ddot{x}_i^A = - \frac{\partial}{\partial x_i^A} V(x, t), \quad i = 1, 2, 3, A = 1, 2, \dots, N \quad (1.21)$$

możemy dość łatwo doprowadzić je do postaci równań Eulera.

W tym celu zauważmy, iż pojedynczy człon po lewej stronie można zapisać następująco:

$$m^A \ddot{x}_i^A = \frac{d}{dt} m^A \dot{x}_i^A = \frac{d}{dt} \frac{\partial}{\partial \dot{x}_i^A} \sum_{j=1, B=1}^{j=3, B=N} \frac{1}{2} m^B (\dot{x}_i^B)^2 = \frac{d}{dt} \frac{\partial}{\partial \dot{x}_i^A} T \quad (1.22)$$

³ Siły są potencjalne, gdy każda i -ta składowa siły działająca na A -ty punkt materialny może być przedstawiona jako pochodna cząstkowa pewnej funkcji **wszystkich położeń** (i ewentualnie czasu) po i -tej współrzędnej cząstki o numerze A .

gdzie oznaczyliśmy literą T sumę członów znanych Wam (niemal od dziecka) jako energie kinetyczne klasycznej mechaniki.

Równania Newtona to $3N$ równań różniczkowych:

$$\frac{d}{dt} \frac{\partial}{\partial \dot{x}_i^A} (T) = \frac{\partial}{\partial x_i^A} (T - V) \quad (1.23)$$

Potencjał V **nie** zależy od prędkości, a wielkość T **nie** zależy od położenia. Dlatego możemy **bezkaranie** do pierwszego nawiasu **dopisać** $-V$, a do drugiego T .

$$\frac{d}{dt} \frac{\partial}{\partial \dot{x}_i^A} (T - V) = \frac{\partial}{\partial x_i^A} (T - V) \quad (1.24)$$

Teraz widzimy, że są to równania Eulera dla funkcjonału:

$$\int (T - V) dt \quad (1.25)$$

W tym kontekście równania wariacyjne nazywają się równaniami Eulera-Lagrange'a. Funkcja podcałkowa $T - V$ nazywa się funkcją Lagrange'a, lub lagranżianem. Oznaczana jest ona zazwyczaj literą L .

$$L = T - V \quad (1.26)$$

Funkcjonał, który dla ruchu rzeczywistego przyjmuje wartość ekstremalną nazywa się **działaniem** (i oznacza często literą S):

$$S = \int (T - V) dt \quad (1.27)$$

Warunek ekstremalności (ściślej **stacjonarności**) działania nazywany jest

zasadą najmniejszego⁴ działania

Wypisanie równań E-L we współrzędnych kartezjańskich prowadzi dokładnie do równań Newtona. Gdyby tylko do tego się to wszystko sprowadzało, gra nie warta by była świeczki. Z funkcjonałem (zwanym działaniem zasada najmniejszego działania) można robić różne operacje prowadzące do niesłychanego uproszczenia rachunków w sytuacjach, gdzie próba zastosowania pocziwych równań Newtona doprowadziłaby do niebotycznych trudności.

⁴ Dla czasów t_1 i t_2 dostatecznie bliskich, rozwiązanie stacjonarne, jest zarazem najmniejsze. Jeśli czasy te są wystarczająco odległe, może się zdarzyć, że jedne modyfikacje ruchu rzeczywistego prowadzą do podwyższenia wartości działania, a inne do jej obniżenia. Zawsze jednak **liniowa** część przyrostu działania znika, gdy badamy zmiany wokół trajektorii rzeczywistej.

W pierwszej (piątkowej) części tego wykładu udowodniłem, że równania Newtona we współrzędnych kartezjańskich dla układu N punktów materialnych poddanych działaniu sił **potencjalnych** (w ogólności zewnętrznych i wewnętrznych), określające **rzeczywisty** ruch układu są **identyczne** z równaniami Eulera – Lagrange’a określającymi położenia jako takie funkcje czasu, które minimalizują (a dokładniej zapewniają **stacjonarność**) funkcjonału (zwanego działaniem):

$$\int_{t_1}^{t_2} (T - V) dt \equiv \int_{t_1}^{t_2} L dt = \text{minimum} \quad (1.28)$$

Gdybyśmy całe życie mieli posługiwać się wyłącznie współrzędnymi kartezjańskimi, a ponadto (czy może przede wszystkim) mieli zajmować się układami cząstek podlegających działaniu **zadanych** sił i niczemu więcej to znaczy nie podlegających działaniu **więzów**), powyższa obserwacja byłaby może ciekawostką (może pozwoliłaby na ustalenie związków mechaniki Newtona z głębszymi teoriami – takimi jak mechanika kwantowa, czy teoria względności), ale nie oznaczałaby żadnego postępu w wyznaczaniu ruchu.

Wszak po to, by wyznaczyć rozwiązanie zagadnienia na stacjonarność funkcjonału, musimy wypisać równania różniczkowe (Eulera - Lagrange’a) a te nie są niczym innym jak równaniami Newtona. Gdybyśmy o zasadzie najmniejszego działania nawet nie słyszeli (a jedynie o równaniach Newtona), wypisalibyśmy do rozwiązywania te same równania.

Sformułowanie wariacyjne praw Newtona pozwala niezwykle **łatwo** przejść do **dowolnych współrzędnych**, a także do układów z **więzami**.

Cóż to są ”dowolne” współrzędne. Z praktyki zeszłorocznej znacie jeden przypadek **wygodnych** współrzędnych, mianowicie biegunowych. Współrzędne biegunowe, to inny sposób parametryzowania położenia punktu w przestrzeni (dwuwymiarowej). Istotą tych współrzędnych są równania definiujące:

$$\begin{aligned} x &= r \cos \varphi, \\ y &= r \sin \varphi. \end{aligned} \quad (1.29)$$

wyrażające jednoznacznie położenie kartezjańskie, jeśli znamy w danej chwili wartości r i φ . Opis ruchu przez podanie $r(t)$ i $\varphi(t)$ jest równie pełny, równie skuteczny jak podanie $x(t)$ i $y(t)$.

Jest jasne, że **dla konkretnego ruchu** funkcje $r(t)$ i $\varphi(t)$ są zupełnie różne od funkcji $x(t)$ i $y(t)$. Różne są więc z pewnością (w sensie czysto rachunkowym, analitycznym) równania różniczkowe spełniane przez współrzędne biegunowe i współrzędne kartezjańskie. Ten sam ruch ma zupełnie inny opis analityczny.

Jest rzeczą ważną uświadomić sobie, że wśród tych różnych opisów (i różnych równań) jedne mogą być dużo **wygodniejsze**, dużo łatwiejsze do znalezienia niż inne. Jednak to, czego nauczył nas Newton, pozwala „od ręki” wypisać tylko równania we współrzędnych kartezjańskich. Przeliczenie równań kartezjańskich do innych współrzędnych (jakie wydają się nam stosowniejsze dla danego problemu) wymaga przeliczenia drugich pochodnych współrzędnych kartezjańskich (wyrażonych wzorami takimi jak (1.29)) i wyrażenia ich przez pierwsze i drugie pochodne nowych współrzędnych, jako funkcji czasu. Pół biedy, gdy są dwie współrzędne kartezjańskie i dwie nowe (biegunowe). Ale gdy współrzędnych jest f , pierwsza pochodna współrzędnej kartezjańskiej:

$$\dot{x}_i^A = x_i^A(q_1, q_2, q_3, \dots, q_f, t) \quad (1.30)$$

to już będzie $f+1$ członów

$$\dot{x}_i^A = \sum_j \frac{\partial x_i^A(q_1, q_2, q_3, \dots, q_f, t)}{\partial q_j} \dot{q}_j + \frac{\partial x_i^A(q_1, q_2, q_3, \dots, q_f, t)}{\partial t} \quad (1.31)$$

a kolejne różniczkowanie wyprodukuje w jednym tylko równaniu f członów z drugimi pochodnymi nowych współrzędnych po czasie, i $(f+1)^2$ członów z niższymi pochodnymi.

W zasadzie wariacyjnej nie ma **wielu** równań, ani drugich pochodnych. Jest jeden skalar – lagranżian – zależny od współrzędnych i tylko od pierwszych pochodnych po czasie. Jeżeli do lagranżianu podstawimy współrzędne kartezjańskie wyrażone wzorami (1.28) i prędkości dane wzorami (1.29), wartośćdziałania od czasu t_1 do t_2 **dla tego samego** ruchu porównawczego będzie **tą samą liczbą**, niezależnie od tego czy obliczamy całkę z użyciem współrzędnych kartezjańskich:

$$\int_{t_1}^{t_2} L(x(t), \dot{x}(t), t) dt ,$$

czy całkę

$$\int_{t_1}^{t_2} L(x(q(t), t), \dot{x}(q(t), \dot{q}(t), t), t) dt \equiv \int_{t_1}^{t_2} \tilde{L}(q(t), \dot{q}(t), t) dt \quad (1.32)$$

z użyciem nowych współrzędnych.

Wartość tej całki dla ruchu **rzeczywistego** jest najmniejsza. **Nie ma znaczenia jak ją obliczamy**. Zależności $q(t)$ opisujące ruch rzeczywisty **minimalizują** całkę (1.32). To jasne jak Słońce! Oczywista oczywistość.

Jeżeli tak, to równania różniczkowe dla zależności $q(t)$ są **po prostu** równaniami Eulera – Lagrange’a dla nowego funkcjonału. (nowy funkcjonał jest liczbowo równy staremu, ale analityczna zależność $\tilde{L}$ od q (i $\dot{q}$) jest **inna** niż starego L od x (i $\dot{x}$). Często będzie to prostsza zależność, a tym samym równania łatwiejsze do rozwiązania.

Zasada najmniejszego działania pozwala na **oszałamiająco** łatwą zamianę zmiennych!

Trzeba podkreślić, że nowe „współrzędne” q , to nie są tylko, zwyczajnie, nowe współrzędne w trójwymiarowej przestrzeni zastępujące współrzędne kartezjańskie wszystkich punktów. Współrzędne q , które zwą się **współrzędnymi uogólnionymi** to zbiór niemal dowolnie wybranych f wielkości wyznaczonych przez konfigurację naszych punktów mających tę własność, że, z kolei, znajomość wszystkich f wielkości q (w danej chwili) określa **jednoznacznie** położenia wszystkich punktów (czyli konfigurację) układu.

Dla pojedynczego punktu w polu centralnym wygodnymi współrzędnymi określającymi położenie naszego punktu (na płaszczyźnie) będą współrzędne biegunowe. Ale, na przykład, dla dwóch punktów materialnych na płaszczyźnie, naturalnymi 4 wielkościami będą dwie współrzędne kartezjańskie środka masy, oraz odległość i azymut wektora względnego położenia ciał.

Możliwymi współrzędnymi do opisu położenia ciała w przestrzeni mogą też być współrzędne **względem obracającego** się układu współrzędnych. Po przeliczeniu lagranżianu (co jest czynnością niezbyt żmudną i nie wymagającą szczególnej pomysłowości), wypisujemy równania E-L i wszystkie siły Coriolisa czy odśrodkowe, jak żywe, wyskakują z formalizmu.

Cząstka w polu centralnym.

Korzystając z (1.28) mamy:

$$\begin{aligned}\dot{x} &= \dot{r} \cos \varphi - r\dot{\varphi} \sin \varphi \\ \dot{y} &= \dot{r} \sin \varphi + r\dot{\varphi} \cos \varphi\end{aligned}\tag{1.33}$$

Dla kwadratu prędkości daje to:

$$v^2 = \dot{x}^2 + \dot{y}^2 = (\dot{r} \cos \varphi - r\dot{\varphi} \sin \varphi)^2 + (\dot{r} \sin \varphi + r\dot{\varphi} \cos \varphi)^2 = \dot{r}^2 + r^2 \dot{\varphi}^2,\tag{1.34}$$

a dla całego lagranżianu:

$$\tilde{L}(r, \varphi, \dot{r}, \dot{\varphi}) = T - V = \frac{m}{2}(\dot{r}^2 + r^2 \dot{\varphi}^2) - V(r)\tag{1.35}$$

Gdy się ma lagranżian, przed przystąpieniem do wypisania równań E-L, należy mu się przyjrzeć, czy nie ma w nim szczególnych symetrii, pozwalających wypisać prawa zachowania, czyli równania pierwszego rzędu (całki układu równań E-L).

W naszym przypadku, nowy Lagranżian **nie zawiera**⁵ zmiennej ϕ . To przyjemna wiadomość. Równanie Lagrange’a odpowiadające tej zmiennej ma 0 po prawej stronie, więc wielkość:

$$\frac{\partial \tilde{L}}{\partial \dot{\phi}} = mr^2 \dot{\phi} = J = \text{constans.} \quad (1.36)$$

Oznaczenie J jest tradycyjne. W procesie wprowadzania 4 stałych (dwa równania 2-go rzędu) zrobiliśmy już krok do przodu. Mamy J .

Lagranżian nie ma też zależności od czasu. Jest więc i druga stała ruchu:

$$\begin{aligned} \dot{r} \frac{\partial \tilde{L}}{\partial \dot{r}} + \dot{\phi} \frac{\partial \tilde{L}}{\partial \dot{\phi}} - \tilde{L} = \\ = m\dot{r}^2 + mr^2 \dot{\phi}^2 - \left(\frac{m\dot{r}^2 + mr^2 \dot{\phi}^2}{2} - V(r) \right) = \frac{m\dot{r}^2 + mr^2 \dot{\phi}^2}{2} + V(r) = E \end{aligned} \quad (1.37)$$

Rozpoznajemy **energię**. Samograj! (Dalej, przez rozdzielanie zmiennych, sprowadzamy problem do obliczenia zwyczajnej całki – elementarnej, bądź nie, w zależności od postaci potencjału).

Przykład jest fundamentalny i bardzo prosty. Dlatego nic szczególnie nowego nie mogliśmy się tu spodziewać. Zwracam jednak uwagę na łatwość i naturalność postępowania. Ani iloczyn wektorowy, ani jakieś szczególne rysunki nie były nam potrzebne. Lagrange, który odkrył powyższy sposób wyznaczania równań w dowolnych współrzędnych, (bez związku z zasadą wariacyjną, a przez dość żmudne przeliczanie wychodząc od równań Newtona), szczycił się tym, że jego podręcznik „Mechaniki analitycznej” nie zawiera ani jednego rysunku!!!

Dwie cząstki.

Przyjmujemy iż potencjał oddziaływania zależy jedynie od odległości. Zatem:

$$L = \frac{m_1}{2} \dot{\vec{r}}_1^2 + \frac{m_2}{2} \dot{\vec{r}}_2^2 - V(|\vec{r}_1 - \vec{r}_2|) \quad (1.38)$$

Całkę energii widzimy. I nic poza tym, gdyż potencjał zależy od wszystkich sześciu współrzędnych: $|\vec{r}_1 - \vec{r}_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$.

Warto zmienić zmienne! Czyli wprowadzić inne „współrzędne”, czyli innych 6 liczb pozwalających wyliczyć stare 6 współrzędnych $(x_1, y_1, z_1, x_2, y_2, z_2)$. Trzy spośród nich to powinny być współrzędne wektora położenia względnego:

⁵ Zauważmy, że we współrzędnych kartezjańskich $L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - V(\sqrt{x^2 + y^2})$ zależy, niestety, od obu współrzędnych x i y .

$$\vec{r} = \vec{r}_1 - \vec{r}_2 \quad (1.39)$$

a trzy inne to może być dowolna inna kombinacja liniowa tych wektorów:

$$\vec{R} = \alpha \vec{r}_1 + \beta \vec{r}_2 \quad (1.40)$$

Łatwo rozwiązać powyższe dwa równania definiujące nowe współrzędne $\vec{r}$ i $\vec{R}$ względem $\vec{r}_1$ i $\vec{r}_2$:

$$\vec{r}_1 = \frac{\vec{R} + \beta \vec{r}}{\alpha + \beta} \quad (1.41)$$

$$\vec{r}_2 = \frac{\vec{R} - \alpha \vec{r}}{\alpha + \beta} \quad (1.42)$$

Teraz już możemy wyznaczyć stary-nowy⁶ lagranżian:

$$\begin{aligned} \tilde{L} &= \frac{m_1}{2} \left(\frac{\dot{\vec{R}} + \beta \dot{\vec{r}}}{\alpha + \beta} \right)^2 + \frac{m_2}{2} \left(\frac{\dot{\vec{R}} - \alpha \dot{\vec{r}}}{\alpha + \beta} \right)^2 - V(|\vec{r}|) = \\ &= \frac{m_1 + m_2}{2} \left(\frac{\dot{\vec{R}}}{\alpha + \beta} \right)^2 + \frac{m_1 \beta^2 + m_2 \alpha^2}{2} \left(\frac{\dot{\vec{r}}}{\alpha + \beta} \right)^2 + (\beta m_1 - \alpha m_2) \frac{\dot{\vec{R}} \cdot \dot{\vec{r}}}{(\alpha + \beta)^2} - V(|\vec{r}|) \end{aligned} \quad (1.43)$$

Na razie nie jesteśmy do końca usatysfakcjonowani rezultatem. Wprowadźmy Lagranżian przestał zależeć od trzech współrzędnych będących składowymi wektora $\vec{R}$ (co sygnalizuje 3 prawa zachowania), ale jego zależność od iloczynu skalarnego prędkości „sprzęga” równania na $\vec{R}$ z równaniami na $\vec{r}$. Członu tego możemy się pozbyć, bo przecież na razie nie ustaliliśmy wartości parametrów α i β . Wzory nasze odpowiadają rodzinie nowych współrzędnych. Nam wystarczy jeden członek tej rodziny, byle wygodny.

W sposób oczywisty nasuwa się wybór:

$$\begin{aligned} \beta m_1 &= \alpha m_2 \\ \alpha + \beta &= 1 \end{aligned} \quad (1.44)$$

czyli:

$$\begin{aligned} \alpha &= \frac{m_1}{m_1 + m_2} \\ \beta &= \frac{m_2}{m_1 + m_2} \end{aligned} \quad (1.45)$$

Ostatecznie mamy:

$$\vec{R} = \frac{m_1 \vec{r}_1 + m_2 \vec{r}_2}{m_1 + m_2} \quad (1.46)$$

⁶ Lagranżian jest „stary”, bo ma starą wartość dla tej samej fizycznej konfiguracji układu, ale „nowy” bo funkcyjnie, jako funkcja nowych współrzędnych jest inny (ma na przykład inne współczynniki).

$$\tilde{L} = \frac{m_1 + m_2}{2} \dot{\vec{R}}^2 + \frac{m_1 m_2}{2(m_1 + m_2)} \dot{\vec{r}}^2 - V(|\vec{r}|) \quad (1.47)$$

Przekonaliśmy się, że obok współrzędnych wektora względnego, drugą naturalną trójką współrzędnych są składowe **wektora środka masy**. Nowy lagranżian jest sumą dwóch części zawierających rozłączne grupy zmiennych. Oznacza to iż układ 6-ciu równań ruchu rozpada się na dwa układy po 3 równania z trzema zmiennymi. Pierwsza część jest identyczna z lagranżianem fikcyjnej cząstki swobodnej (o masie równej sumie mas rzeczywistych cząstek), druga z lagranżianem innej quasicząstki o masie zredukowanej poruszającej się w centralnym polu sił.