

Rozdział 6

Rachunek wariacyjny

6.1 Typowy problem wariacyjny

Początki rachunku wariacyjnego to koniec XVII wieku. W 1696 r., Jan Bernoulli w *Acta Eruditorum*¹ wezwał „najtęższych matematyków świata” do rozwiązania *zagadnienia brachistochrony*: znalezienia równania krzywej, po której porusza się punkt materialny pod wpływem sił ciężkości, a która ma własność zminimalizowania czasu podróży pomiędzy dwoma, dowolnie zadanymi, punktami². Problem ten dobrze ilustruje podstawowe zagadnienie rachunku wariacyjnego, jakim jest znalezienie ekstremum (praktycznie zawsze to będzie minimum) *funkcjonału* – wielkości wyrażonej za pomocą całki, której wartość zależy od *drogi całkowania*, a więc postaci krzywej łączącej dwie konkretne wartości zmiennej całkowania.

W problemie brachistochrony zmienną x – *zmienną niezależną* – jest odległość „w poziomie” pomiędzy położeniami punktu; położenie „w pionie” $y = y(x)$ to *zmienna zależna*. Dwoma punktami, pomiędzy którymi przebiega droga całkowania,

¹Był to pierwszy niemiecki periodyk naukowy. Ukazywał się on raz w miesiącu i był wydawany w Lipsku, w latach 1682–1782, przez Otto Menke i jego syna Johanna. Ten naukowy miesięcznik wydawany był po łacinie, a publikowali w nim najlepsi uczeni, żyjący w tym – jakże produktywnym – stuleciu.

²Brachistochroną jest – jak wiemy i zaraz zobaczymy – cykloida. Jest ona zarazem *izochroną* albo *tautochroną* – czas podróży będzie taki sam dla różnych (leżących na różnych wysokościach) punktów startowych. Nazwy wywodzą się z języka greckiego: *brakhus* – krótki; *tauto* – taki sam; *isos* – równy, ten sam; *chronos* – czas. Empirycznie problem ten był już znany Christianowi Huyghensowi, który wykorzystywał izochronizm cykloidy do konstrukcji zegarów. Problem brachistochrony rozwiązali: Newton, Leibniz, de l’Hospital (n.b. uczeń Jana B.) oraz „wielki brat Jana”, Jakub Bernoulli. Z ciekawostek: Jakub znał prawidłową odpowiedź już od pewnego czasu, i zachodzą podejrzenia, że jego (młodszy) brat . . . zapoznał się z nią przed ogłoszeniem konkursu. Leibniz nie tylko rozwiązał problem, ale podał (prawidłowo!) nazwiska wszystkich matematyków, którzy z nim się uporają. Newton, będący już w nienajlepszej kondycji psychicznej, po otrzymaniu wyzwania zamknął się na cały wieczór w swojej pracowni i wyszedł z niej o 4. rano z gotowym rozwiązaniem.

są punkt „startu” (x_s, y_s) i punkt „mety” (x_m, y_m) – *więzy* narzucone na ruch

$$y_s = y(x_s), \quad y_m = y(x_m). \quad (6.1)$$

Nasze zadanie to znalezienie równania krzywej $y = y(x)$, które musi spełniać *więzy* (6.1) i minimalizuje całkowity czas T podróży, który możemy przedstawić w postaci całki

$$T = \int_{x_s, y_s}^{x_m, y_m} dt = \int_{x_s, y_s}^{x_m, y_m} \frac{ds}{v}, \quad (6.2)$$

gdzie ds to element drogi, a v – prędkość podróżującego ciała. Wartość tej ostatniej obliczamy z prawa zachowania energii

$$v = v(y) = \sqrt{2g(y_s - y)}; \quad (6.3)$$

z kolei

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \sqrt{1 + y_x^2} dx, \quad (6.4)$$

gdzie przez y_x będziemy oznaczać dy/dx . Po podstawieniu z (6.3) i (6.4) do (6.1) mamy³

$$T = \int_{x_m}^{x_s} \frac{\sqrt{1 + y_x^2}}{\sqrt{2g(y_s - y)}} dx. \quad (6.5)$$

Określenie krzywej najszybszego spadku, to wybranie, spośród nieskończenie wielu możliwych trajektorii łączących punkt (y_s, x_s) („start”) z punktem (x_m, y_m) („meta”) tej, która minimalizuje wyrażenie (6.5).

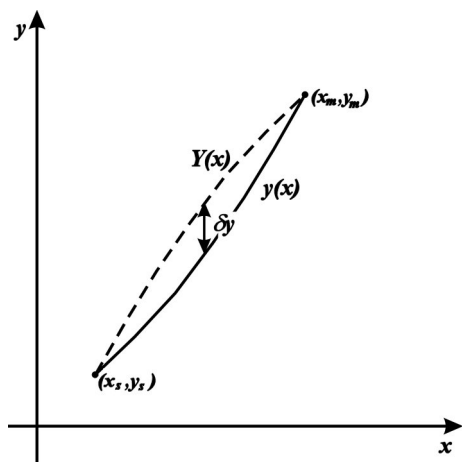
6.2 Rachunek wariacyjny – jedna zmienna niezależna i jedna zmienna zależna

Sformułowany w poprzednim podrozdziale problem ilustruje najprostszy przypadek rachunku wariacyjnego, kiedy mamy do czynienia z jedną funkcją – *jedną zmienną zależną* – $y = y(x)$, będącą funkcją (tylko!) *jednej zmiennej niezależnej* x . Występująca pod znakiem całki funkcja f zależy od obu zmiennych, zależnej i niezależnej, a także od pochodnej $dy/dx \equiv y_x$. Ogólne sformułowanie problemu wariacyjnego sprowadza się do znalezienia $y = y(x)$, dla której funkcjonal \mathcal{J} , określony jako

$$\mathcal{J} = \int_{x_1}^{x_2} f\left(y(x), \frac{dy}{dx}, x\right) dx, \quad (6.6)$$

przyjmuje wartość ekstremalną. Taką drogę całkowania oznaczamy $y(x)$. W kontekście rozważań mechanicznych możemy nazwać ją *trajektorią rzeczywistą*. Dla każdej wartości zmiennej x z przedziału (x_1, x_2) zdefiniujemy *trajektorię porównawczą* $Y(x)$, przebiegającą w bezpośrednim sąsiedztwie $y(x)$.

³ „Wymiana” granic całkowania to konsekwencja antyrównoległości *wektorów* ds i v .



Rysunek 6.1: Trajektoria rzeczywista $y(x)$, minimalizująca wartość funkcjonału \mathcal{J} i trajektoria porównawcza $Y(x)$, obie przechodzące przez $y(x_1) = Y(x_1) \equiv y_1$ oraz $y(x_2) = Y(x_2) \equiv y_2$

Podkreślmy, że *wszystkie* trajektorie przechodzą przez narzucone punkty (x_1, y_1) i (x_2, y_2) – por. rysunek 6.1.

Odchylenie trajektorii porównawczej od rzeczywistej w punkcie x nazywamy *wariacją funkcji y* i oznaczamy

$$\delta y = \delta y(x) = Y(x) - y(x). \quad (6.7)$$

Ze względu na narzucone wartości $y_1 = y(x_1) = Y_1(x_1)$ oraz $y_2 = y(x_2) = Y_2(x_2)$ dla końców przedziału mamy

$$\delta y(x_1) = \delta y(x_2) = 0. \quad (6.8)$$

Podobnie możemy wprowadzić, różniczkując obie strony (6.7), wariację *pocho-
dnej funkcji $y(x)$* dy/dx

$$\delta y_x = \frac{dY(x)}{dx} - \frac{dy(x)}{dx} = \frac{d}{dx} \delta y. \quad (6.9)$$

Dla funkcji f , będącej w ogólnym przypadku funkcją y i y_x , wariacja δf będzie się wyrażała przez wariacje δy i δy_x ⁴

$$\delta f \equiv \delta f(x) = f\left(Y(x), \frac{dY}{dx}, x\right) - f\left(y(x), \frac{dy}{dx}, x\right) = \frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y_x} \delta y_x. \quad (6.10)$$

Konsekwentnie *wariacja funkcjonału* będzie określona jako

$$\begin{aligned} \delta \mathcal{J} &= \mathcal{J}(Y, Y_x, x) - \mathcal{J}(y, y_x, x) = \int_{x_1}^{x_2} \left[f\left(Y(x), \frac{dY}{dx}, x\right) - f\left(y(x), \frac{dy}{dx}, x\right) \right] dx \\ &= \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y_x} \frac{d}{dx} \delta y \right] dx. \end{aligned} \quad (6.11)$$

Dla trajektorii $y(x)$ realizującej minimum \mathcal{J} , wariacja wielkości \mathcal{J} musi być równa zero. Mamy tu analogiczną sytuację jak w przypadku „zwykłego” warunku ekstremum dla funkcji $g = g(x)$ – aby $g(x)$ osiągało ekstremum musi być $dg = 0$. Warunkiem dla trajektorii rzeczywistej będzie więc

$$\delta \mathcal{J} = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} \delta y + \frac{\partial f}{\partial y_x} \frac{d}{dx} \delta y \right] dx = 0. \quad (6.12)$$

⁴W wyrażeniu na *wariację funkcji f* nie występuje pochodna cząstkowa $\partial f/\partial x$ – zmiany funkcjonału wynikają bowiem z odstępstwa – *przy ustalonej wartości x* – trajektorii porównawczej od rzeczywistej. Inaczej mówiąc: na zmianę funkcjonału mają wpływ wariacje funkcji δy i jej pochodnej δy_x , ale „wariacja” $\delta x = 0$.

Drugi człon w funkcji podcałkowej „zaprasza” do całkowania przez części. Mamy

$$\begin{aligned} \int_{x_1}^{x_2} \frac{\partial f}{\partial y_x} \frac{d}{dx} \delta y \, dx &= \int_{x_1}^{x_2} \left[\frac{d}{dx} \left(\frac{\partial f}{\partial y_x} \delta y \right) - \delta y \frac{d}{dx} \frac{\partial f}{\partial y_x} \right] dx \\ &= \left. \frac{\partial f}{\partial y_x} \delta y \right|_{x=x_1}^{x=x_2} - \int_{x_1}^{x_2} \delta y \frac{d}{dx} \frac{\partial f}{\partial y_x} dx. \end{aligned} \quad (6.13)$$

W tym momencie wykorzystujemy „warunek więzów” (6.8) – z całki (6.13) pozostaje tylko drugi wyraz i całka wyjściowa (6.12) przybiera postać

$$\delta \mathcal{J} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y_x} \right) \delta y \, dx; \quad (6.14)$$

warunek zerowania się jej dostarcza – z uwagi na *zupełną dowolność w wyborze naszych wariacji* δy – równania na szukaną trajektorię rzeczywistą $y(x)$ ⁵.

$$\left(\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y_x} \right) \delta y(x) = 0, \quad (6.15)$$

albo

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y_x} = 0.} \quad (6.16)$$

Równanie (6.16) nosi nazwę równania Eulera-Lagrange’a (E-L)⁶. Występuje ono w różnych postaciach, które mają swoje zastosowanie w pewnych, specyficznych przypadkach. Na przykład obliczając

$$\frac{df}{dx} = \frac{\partial f}{\partial y} y_x + \frac{\partial f}{\partial y_x} y_{xx} + \frac{\partial f}{\partial x}$$

oraz

$$\frac{d}{dx} \left(y_x \frac{\partial f}{\partial y_x} \right) = y_{xx} \frac{\partial f}{\partial y_x} + y_x \frac{d}{dx} \left(\frac{\partial f}{\partial y_x} \right),$$

zweryfikujemy łatwo, że

$$\frac{\partial f}{\partial x} - \frac{d}{dx} \left(f(x) - y_x \frac{\partial f}{\partial y_x} \right) = -y_x \left[\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y_x} \right],$$

⁵Zerowanie się całki nie musi wynikać z zerowania się funkcji podcałkowej. Ale *dowolność* wariacji y -a pozwala stwierdzić, że aby całka (6.14) była równa zero, to zero *musi być równa* funkcja podcałkowa.

⁶Czasami mówi się po prostu o równaniu Eulera. Zarówno Euler, jak i Lagrange – dwaj najwięksi matematycy 18. wieku – pracowali w Królewskiej (Berlińskiej) Akademii Nauk (u Fryderyka Wielkiego), chociaż ani jeden, ani drugi nie pozostali tam zbyt długo. Euler powrócił – na zaproszenie Katarzyny Wielkiej – do Petersburga, Lagrange – z pochodzenia Włoch – przyjął zaproszenie Ludwika XVI do Francji, gdzie przyszło mu przeżyć czasy Wielkiego Terroru.

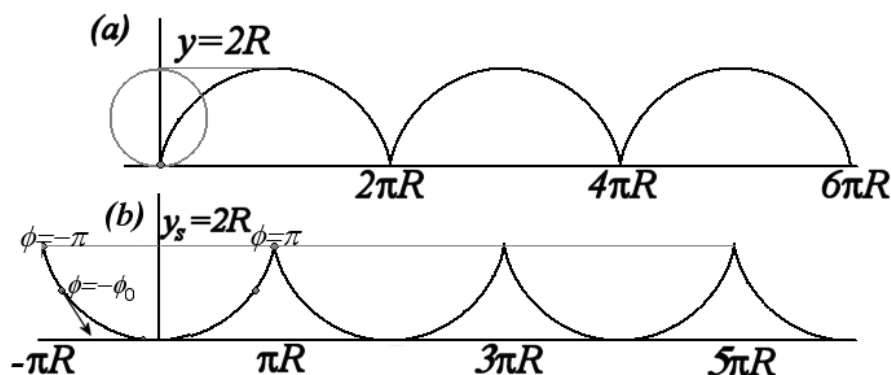
a jeżeli tak to alternatywna postać równania Eulera-Lagrange’a to właśnie

$$\frac{\partial f}{\partial x} - \frac{d}{dx} \left(f(x) - y_x \frac{\partial f}{\partial y_x} \right) = 0. \quad (6.17)$$

Jest ona szczególnie przydatna, jeżeli funkcja $f(x)$ nie zależy *explicite* od x , gdyż wtedy równanie E-L przybiera już całkiem prostą postać

$$f(x) - y_x \frac{\partial f}{\partial y_x} = \text{pewna stała}. \quad (6.18)$$

Nie zapominajmy, że w równaniu (6.16) lub (6.18) dysponujemy jawną postacią funkcji f , natomiast szukaną funkcją jest „rzeczywista trajektoria” $y(x)$.



Rysunek 6.2: Cykloida, krzywa jaką zakresła punkt obwodu koła, toczącego się ze stałą prędkością kątową po poziomej prostej(a); cykloida odwrócona, brachistochrona siły ciężkości(b)

Jako przykład zastosowania równania Eulera-Lagrange’a rozwiążemy do końca problem brachistochrony, równanie (6.5), w którym położymy $x_m = y_m = 0$ [„meta” w początku układu – por. rysunek 6.2(b)].

$$T = \int_0^{x_s} \frac{\sqrt{1 + y_x^2}}{\sqrt{2g(y_s - y)}} dx. \quad (6.19)$$

Występująca w nim funkcja podcałkowa to

$$f = f(y, y_x) = \sqrt{\frac{1 + y_x^2}{2g(y_s - y)}} \quad (6.20)$$

nie zależy od x , a więc stosujemy uproszczoną postać równania Eulera (6.18). Podstawiamy do niej $f(y, y_x)$ z równania (6.20) i po prostych przekształceniach mamy

$$\sqrt{(y_s - y)(1 + y_x^2)} = \text{stała} \equiv C. \quad (6.21)$$

„Rozwikłanie” tego równania ze względu na pochodną dy/dx prowadzi do

$$\frac{dy}{dx} = \sqrt{\frac{C^2 - y_s + y}{y_s - y}}. \quad (6.22)$$

W dalszych rachunkach pomożemy sobie ...znajomością końcowego wyniku. Brachistochroną ma być „odwrócona” (wypukłością ku dołowi) cykloida, a więc krzywa zakreślana przez punkt obwodu koła o promieniu R , toczącego się po prostej ze stałą prędkością kątową ω . Parametryczne równania cykloidy (koło i punkt na jego obwodzie startują z początku układu współrzędnych) to

$$x(\phi) = R(\phi - \sin \phi) \quad (6.23)$$

$$y(\phi) = R(1 - \cos \phi), \quad (6.24)$$

gdzie parametr $\phi = \omega t$ (t to czas). Taką cykloidę widzimy na rysunku 6.2(a).

Aby nasze obliczenia były jak najprostsze, wybierzmy stałą C^2 w równaniu (6.22) równą y_s ⁷. Równanie to przybiera wówczas postać

$$\frac{dy}{dx} = \sqrt{\frac{y}{y_s - y}}. \quad (6.25)$$

Ponieważ wiemy, że jego rozwiązaniem ma być cykloida, zdajemy sobie sprawę, że trudno byłoby liczyć na uzyskanie równania w postaci uwikłanej $y = y(x)$; zamiast tego spróbujmy dokonać *a priori* parametryzacji równania, „inspirując się” wzorem (6.24) i pisząc

$$y = y(\phi) = R(1 - \cos \phi), \quad (6.26)$$

gdzie kładziemy – zgodnie z rysunkiem 6.2 – $R = y_s/2$ (współrzędna y -owa punktu startu to najwyższe z możliwych położenia punktu na kole o średnicy $2R$). Mamy więc

$$y = y(\phi) = \frac{y_s}{2}(1 - \cos \phi); \quad (6.27)$$

podstawiając z tego równania do (6.25), dostaniemy

$$\frac{dy}{dx} = \sqrt{\frac{\frac{y_s}{2} - \frac{y_s}{2} \cos \phi}{y_s - \frac{y_s}{2} + \frac{y_s}{2} \cos \phi}} = \sqrt{\frac{1 - \cos \phi}{1 + \cos \phi}}, \quad (6.28)$$

⁷Warto to przemyśleć „do końca”. Z pewnością możemy wybrać różnicę $C^2 - y_s$ jako pewną nową – różną od zera – stałą, a to oznacza: (1) „inny początek” osi y -ów oraz (2) „inną wartość” współrzędnej y_s (inny punkt startu) w mianowniku wzoru (6.22). Jak to potwierdzą uzyskane wyniki, wybór $C^2 = y_s$, a także wybór odpowiedniej stałej całkowania w równaniu na drugą zmienną: $x = x(\phi)$ [por. równanie (6.30)], odpowiadają sytuacji, kiedy punkt $(0, 0)$ stanowi *rzeczywiście* „metę” dla ześlizgującego się po krzywej ciała i jednocześnie jest to punkt przegięcia cykloidy, dla którego wartość pochodnej $y_x = 0$.

albo – uwzględniając (6.27) –

$$\begin{aligned} dx &= \sqrt{\frac{1 + \cos \phi}{1 - \cos \phi}} dy = \frac{y_s}{2} \sqrt{\frac{1 + \cos \phi}{1 - \cos \phi}} \sin \phi d\phi \\ &= \dots = \frac{y_s}{2} (1 + \cos \phi) d\phi. \end{aligned} \quad (6.29)$$

Rozwiązaniem tego równania (z dokładnością do stałej \mathcal{B}) będzie

$$x = \frac{y_s}{2} (\phi + \sin \phi) + \mathcal{B}. \quad (6.30)$$

Stałą \mathcal{B} musimy jednak położyć jako równą zero, jeżeli chcemy aby nasza krzywa przechodziła przez punkt $(0, 0)$. Równania brachistochrony przybierają w tej sytuacji postać

$$x(\phi) = R(\phi + \sin \phi) \quad (6.31)$$

$$y(\phi) = R(1 - \cos \phi), \quad (6.32)$$

i przedstawiają cykloidę, obróconą wypukłością w dół [por. rysunek 6.2(b)]. Pozostawiamy Czytelnikowi sprawdzenie, że (1) „przesuwając” argument kątowy o π : $\phi \rightarrow \pi + \phi$ w obu wzorach; (2) dodając do współrzędnej x -owej stałą $-\pi R$; i (3) „odbijając” naszą trajektorię w osi $0x$ oraz przesuwając ją w górę o $2R$: $y \rightarrow -y + 2R$, otrzymamy równania (6.23) i (6.24) opisujące „normalną” krzywą cykloidy, jaką widzimy na rysunku 6.2(a).

Cykloida jest krzywą najszybszej podróży, ale warto policzyć ten czas, aby stwierdzić naocznie, że jest on niezależny od współrzędnych (x_s, y_s) punktu startowego. Zgodnie z (6.5) (pamiętajmy: $x_m = y_m = 0$)

$$T = \frac{1}{\sqrt{2g}} \int_{x_s}^0 \frac{\sqrt{1 + y_x^2}}{\sqrt{y_s - y}} dx. \quad (6.33)$$

Teraz za y_x podstawiamy z (6.28), za $y - z$ (6.26) i wreszcie zamiast całkować względem x podstawiamy z (6.31) $dx = R(1 + \cos \phi) d\phi$. Obliczmy czas podróży dla sytuacji, kiedy punkt startuje z najwyższego położenia – będzie to odpowiadało, zgodnie z (6.32), przyjęciu $y_s = R[1 - \cos(-\pi)] = 2R$ i całkowaniu w granicach [por. rysunek 6.2(b)] $\phi = -\pi$ i $\phi = 0$. Ponieważ [por. (6.27)–(6.29)]

$$\sqrt{1 + y_x^2} = \sqrt{\frac{2}{1 + \cos \phi}}; \quad \frac{\sqrt{1 + y_x^2}}{\sqrt{y_s - y}} = \sqrt{\frac{2}{R}} \frac{1}{1 + \cos \phi}, \quad (6.34)$$

całka (6.5) przyjmuje prostą postać

$$T = \sqrt{\frac{R}{g}} \int_{-\pi}^0 d\phi = \pi \sqrt{\frac{R}{g}}. \quad (6.35)$$

To było dość łatwe. Policzenie czasu podróży dla *dowolnego punktu startu*, a więc dla dowolnej wartości $\phi = -\phi_0$ [por. rysunek 6.2(b)] wymaga nieco więcej zachodu. Przyjmujemy $y_s = R[1 - \cos(-\phi_0)]$ i całkujemy w granicach [por. rysunek 6.2(b)] $\phi = -\phi_0$ i $\phi = 0$. Odpowiednikiem (6.33) będzie

$$\begin{aligned} T &= \frac{1}{\sqrt{2g}} \int_{-\phi_0}^0 \sqrt{\frac{2}{1 + \cos \phi}} \frac{R(1 + \cos \phi) d\phi}{\sqrt{R(1 - \cos \phi_0) - R(1 - \cos \phi)}} \\ &= \sqrt{\frac{R}{g}} \int_{-\phi_0}^0 \sqrt{\frac{1 + \cos \phi}{\cos \phi - \cos \phi_0}} d\phi \\ &= \sqrt{\frac{R}{g}} \int_{-\phi_0}^0 \frac{\sqrt{2} \cos \frac{\phi}{2}}{\sqrt{\cos^2 \phi/2 - \sin^2 \phi/2 - \cos^2 \phi_0/2 + \sin^2 \phi_0/2}} d\phi \\ &= \sqrt{\frac{R}{g}} \int_{-\phi_0}^0 \frac{\cos \frac{\phi}{2} d\phi}{\sqrt{\sin^2 \phi_0/2 - \sin^2 \phi/2}} = \boxed{u = \frac{\sin \phi/2}{\sin \phi_0/2}, \quad du = \frac{\cos \phi/2}{2 \sin \phi_0/2}} \\ &= 2\sqrt{\frac{R}{g}} \int_{-1}^0 \frac{du}{\sqrt{1 - u^2}} = 2\sqrt{\frac{R}{g}} \arcsin u \Big|_{-1}^0 = \pi\sqrt{\frac{R}{g}}. \end{aligned}$$

Rachunki rzeczywiście nieco żmudne, ale satysfakcja ogromna – rzeczywiście, dla *każdego punktu* cykloidy (odpowiadającej pewnej konkretnej wartości parametru R) czas podróży do początku układu będzie taki sam.

Powróćmy jeszcze do pewnych kwestii praktycznych. Rachunki rachunkami, ale zapewne Dociekliwy Czytelnik zadaje sobie pytanie: jak w praktyce zrealizować takie *izochroniczne wahadło cykloidalne*. Odpowiedź na to pytanie dał Christiaan Huygens (tak, ten od zasady Huygensa) i jego współpracownik Salomon Henrikszoon Coster, w roku 1657. To Coster wykonał konstrukcję, według wskazówek Ch. H.

Huygens miał pełną świadomość faktu, że „zwykle” wahadło, którego torem jest łuk okręgu, stanowi kiepski zegar – przy większych wychyleniach jego okres przestaje być stały. W (dużym!) uproszczeniu jego pomysł sprowadzał się do poddania ruchu wahadła pewnym więzom, przez umieszczenie nici wahadła pomiędzy odpowiednio ukształtowanymi „szczękami”, które odpowiednio ograniczają wychylenia.

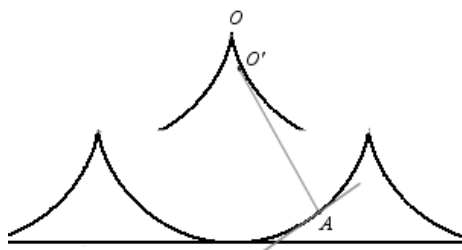
Przy dużych wychyleniach nić kładzie się na szczękach i jej chwilowa długość ulega skróceniu, co automatycznie koryguje wynikającą z większego wychylenia zmianę okresu. Przedstawione to jest na rysunku 6.3. Punkt O' oznacza punkt, przez który przechodzi oś obrotu wahadła, którego „chwilowa” długość $O'A$ jest mniejsza od rzeczywistej – OA .

Łatwo zauważyć, że profil szczęk, to miejsce geometryczne środków okręgów stycznych do trajektorii (matematycznego) wahadła (por. rysunek 6.3), którą jest cykloida. Geometrycznie zagadnienie sprowadza się do konstrukcji *ewoluty* cykloidy (cykloidalna trajektoria jest z kolei jest *ewolwentą* swojej ewoluty).

Stosunkowo proste rachunki opierają się na równaniu

$$\sin \alpha = \sqrt{\frac{y}{2R}}, \quad (6.36)$$

w którym α to kąt jaki tworzy styczna do cykloidy – tej zwykłej z rysunku 6.2(a) – z pionem ($\alpha = \phi/2$) w punkcie o współrzędnej pionowej równej y .



Rysunek 6.3: Cykloidalne szczęki zapewniające ruch po trajektorii cykloidalnej. Punkt O' oznacza „chwilową” oś obrotu wahadła, którego „chwilowa” długość to $O'A$ jest mniejsza od rzeczywistej – OA

Huygens bez trudu wykazał, że ewolucją cykloidy – obróconej wypukłościami w dół – jest *identyczna* cykloida, przesunięta w górę o $2R$, i w bok o πR .

Szczęki powinny być profilami cykloidalnymi ...⁸

Warto może jeszcze zauważyć, że równanie E-L odpowie oczywiście na pytanie: jakie jest równanie krzywej, łączącej dwa punkty (x_1, y_1) oraz (x_2, y_2) na płaszczyźnie, dla której odległość $S(x_1, x_2)$, mierzona właśnie wzdłuż tej krzywej, jest najkrótsza. Wykorzystując równanie (6.4) otrzymamy prosty funk-

cjonal

$$S = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{1 + y_x^2} dx, \quad (6.37)$$

tak więc

$$f = f(y_x) = \sqrt{1 + y_x^2}$$

⁸To brzmi dość prosto, ale detale konstrukcyjne, a także historia pierwszych zegarów izochronicznych nie są ani proste, ani dostatecznie precyzyjnie poznane (por. np. artykuł Alana Emmersona *Things are Seldom What They Seem—Christiaan Huygens, the Pendulum, and the Cycloid* na stronie <http://www.nawcc.org/index.php/watch-a-clock-bulletin/w-a-c-bulletin-addenda> – NAWCC to anagram Narodowego Zrzeszenia Kolekcjonerów Zegarów i Zegarków.).

Dodajmy, że problem konstrukcji izochronicznego, a więc rzetelnego zegara, był wysoce „praktyczny”. Gorączka odkryć geograficznych zapoczątkowana z końcem 15. wieku wywołała zapotrzebowanie na precyzyjny pomiar czasu – dla wyznaczania długości geograficznej. Z końcem 16. wieku problem stał się tak palący, że np. król Hiszpanii Filip II ogłosił konkurs, z niebagatelną nagrodą, na wynalezienie morskiego chronometru. Zainteresowania Huygensa pomiarami czasu nie były więc pozbawione aspektu praktycznego (były to lata kiedy Niderlandy stanowiły część imperium hiszpańskiego). Ale H. nie zdobył nagrody. Upłynęło wiele czasu – nową premię dla wynalazcy ogłosił brytyjski Parlament. O pasjonującej historii konstrukcji pierwszego morskiego chronometru, dokonanej przez angielskiego stolarza i zegarmistrza-samouka Johna Harrisona w drugiej połowie 18. wieku (!) warto przeczytać obszerną notę w Wikipedii (ale nie w polskiej, tylko angielskiej) — http://en.wikipedia.org/wiki/John_Harrison.

i równanie (6.18) daje

$$-\frac{d}{dx} \left(f - y_x \frac{\partial f}{\partial y_x} \right) = -\frac{d}{dx} \left(\sqrt{1 + y_x^2} - \frac{y_x^2}{\sqrt{1 + y_x^2}} \right) = -\frac{d}{dx} \frac{1}{\sqrt{1 + y_x^2}} = 0 \quad (6.38)$$

z oczywistym rozwiązaniem

$$y_x = \frac{dy}{dx} = \text{stała} \equiv a, \quad (6.39)$$

skąd

$$y = ax + b.$$

Dwie stałe całkowania wyznaczmy oczywiście z układu równań

$$\begin{aligned} y_1 = y(x_1) &= ax_1 + b, \\ y_2 = y(x_2) &= ax_2 + b. \end{aligned}$$

6.3 Jedna zmienna niezależna i kilka zmiennych zależnych

Omawiany w poprzednich podrozdziałach problem wariacyjny można poddać różnego rodzaju uogólnieniom, zwiększając liczbę zmiennych zależnych, niezależnych lub zmiennych obu rodzajów. Niewątpliwie, największy sukces rachunku wariacyjnego w zastosowaniach fizycznych, a konkretnie, w mechanice układu punktów materialnych, wiąże się z sytuacją, kiedy mamy do czynienia z jedną zmienną niezależną, czasem t , od którego zależy cały szereg zmiennych zależnych – *współrzędnych uogólnionych* $q_1(t), q_2(t), \dots, q_f(t)$, które opisują ruch układu. Współrzędne te mogą być zwykłymi współrzędnymi, w układzie współrzędnych kartezjańskich, bądź współrzędnymi krzywoliniowych, albo pewnymi funkcjami takich zwykłych współrzędnych, utworzonymi w sposób zapewniający najbardziej „praktyczny” opis układu. Podstawowym kryterium jakie spełnia zupełny zbiór współrzędnych uogólnionych, jest ich liczba, równa liczbie stopni swobody układu f . I tak, dla układu n punktów materialnych, poruszających się w przestrzeni 3-wymiarowej

$$f = 3n - r, \quad (6.40)$$

gdzie r to liczba równań więzów spełnianych przez „zwykłe” współrzędne (np. kartezjańskie) naszych punktów $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$. Na przykład, dwa punkty materialne o współrzędnych $r_1(x_1, y_1, z_1)$ i $r_2(x_2, y_2, z_2)$, których odległość

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

jest stała, możemy opisywać za pomocą pięciu dowolnie wybranych współrzędnych kartezjańskich, wyliczając szóstą z powyższego równania, ale możemy też wprowadzić 3 współrzędne środka masy dwóch punktów i dwa kąty, określające przestrzenną orientację dwójki punktów. Z kolei, dla opisu ruchu wahadła matematycznego (ruch płaski) możemy użyć jako współrzędnej uogólnionej jednej z dwójki współrzędnych kartezjańskich i wyliczać drugą z warunku stałej długości nici wahadła, ale zdecydowanie wygodniejszym będzie użycie kąta ϕ , jaki tworzy nić z pionem.

W kontekście mechanicznym będzie na interesowała minimalizacja funkcjonału

$$\mathcal{J} = \int_{t_1}^{t_2} \mathcal{F}(q_1, q_2, \dots, q_i, \dots, q_f, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_i, \dots, \dot{q}_f, t) dt; \quad \dot{q}_i \equiv \frac{dq_i}{dt}, \quad (6.41)$$

poddanego dodatkowym warunkom, określającym wartości jednoznacznie wartości zmiennych zależnych na krańcach przedziału całkowania $q_1(t_1), \dots, q_f(t_1)$ oraz $q_1(t_2), \dots, q_f(t_2)$. Warunek minimum to zerowanie się wariacji \mathcal{J} , którą wyrazimy poprzez wariację funkcji \mathcal{F} , równą

$$\delta\mathcal{F} = \frac{\partial\mathcal{F}}{\partial q_1} \delta q_1 + \frac{\partial\mathcal{F}}{\partial q_2} \delta q_2 + \dots + \frac{\partial\mathcal{F}}{\partial q_f} \delta q_f + \frac{\partial\mathcal{F}}{\partial \dot{q}_1} \delta \dot{q}_1 + \frac{\partial\mathcal{F}}{\partial \dot{q}_2} \delta \dot{q}_2 + \dots + \frac{\partial\mathcal{F}}{\partial \dot{q}_f} \delta \dot{q}_f. \quad (6.42)$$

Przyrównując wariację (6.41) do zera i stosując identyczne procedury obliczeniowe jak te, które doprowadziły nas do równania (6.13) (całkujemy przez części i korzystamy z zerowania się wszystkich wariacji $q_i(t_1)$ oraz $q_i(t_2)$; $i = 1, \dots, f$), otrzymamy

$$\int_{t_1}^{t_2} \left[\left(\frac{\partial\mathcal{F}}{\partial q_1} - \frac{d}{dt} \frac{\partial\mathcal{F}}{\partial \dot{q}_1} \right) \delta q_1 + \dots + \left(\frac{\partial\mathcal{F}}{\partial q_f} - \frac{d}{dt} \frac{\partial\mathcal{F}}{\partial \dot{q}_f} \right) \delta q_f \right] dt = 0. \quad (6.43)$$

I znowu – ze względu na zupełną dowolność w doborze niezależnych wariacji δq_i – całka (6.43) będzie równa zero, jeżeli każde z wyrażeń w nawiasach będzie, niezależnie, równe zero. Zamiast jednego równania Eulera [przypadek *jednej* zmiennej zależnej, wzór (6.16)] otrzymujemy układ f równań

$$\boxed{\frac{\partial\mathcal{F}}{\partial q_i} - \frac{d}{dt} \frac{\partial\mathcal{F}}{\partial \dot{q}_i} = 0 \quad i = 1, 2, \dots, f.} \quad (6.44)$$

Jeżeli funkcją \mathcal{F} – współrzędnych uogólnionych i czasu – jest lagranżjan (funkcja Lagrange'a) układu L – różnica energii kinetycznej T i potencjalnej V

$$L = T - V \quad (6.45)$$

układu punktów, to równanie

$$\delta L = \delta \int_{t_1}^{t_2} L dt = \delta \int_{t_1}^{t_2} (T - V) dt = 0, \quad (6.46)$$

z którego wynika, że całka względem czasu z lagranżjanu układu musi, w ruchu rzeczywistym, przybierać wartość ekstremalną, nazywa się zasadą (albo: zasadą całkową) Hamiltona. Oparty na niej formalizm mechaniki teoretycznej stanowi nie tylko zdecydowany krok „do przodu” w rozwoju mechaniki newtonowskiej, ale znajduje swoje bliskie paralele w mechanice kwantowej. Dodajmy jeszcze bardzo istotną uwagę: tak sformułowana zasada całkowita odnosi się do przypadku, kiedy siły działające w układzie są zachowawcze, to znaczy takie, które nie zależą *explicite* od czasu t

$$V = V(q_1, q_2, \dots, q_i, \dots, q_f), \quad \frac{\partial V}{\partial t} = 0, \quad \frac{\partial V}{\partial \dot{q}_i} = 0, \quad i = 1, \dots, f. \quad (6.47)$$

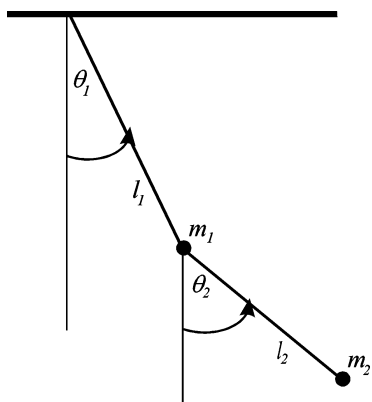
W takim przypadku, równania (6.44) przybierają postać

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial(T - V)}{\partial q_i} - \frac{d}{dt} \frac{\partial(T - V)}{\partial \dot{q}_i} = 0 \quad i = 1, 2, \dots, f, \quad (6.48)$$

a uwzględniając (6.47) mamy

$$\frac{\partial T}{\partial q_i} - \frac{\partial V}{\partial q_i} + \frac{d}{dt} \frac{\partial V}{\partial \dot{q}_i} = 0 \quad \text{albo} \quad \frac{\partial T}{\partial q_i} - \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} = \frac{\partial V}{\partial q_i} \quad i = 1, 2, \dots, f. \quad (6.49)$$

Równania te, określające ruch układu punktów o f stopniach swobody, poddanych działaniu sił zachowawczych, noszą nazwę *równań Lagrange'a drugiego rodzaju*⁹. Stanowią one bardzo „poręczny” formalizm opisu ruchu, szczególnie w przypadku istnienia pewnych ograniczeń (więzów). Takie ograniczenia uwzględnimy dobierając odpowiednio współrzędne uogólnione (opisujące formalnie układ punktów); konkretne postacie tych współrzędnych (ich zależność od czasu) znajdujemy z równań (6.48).



Rysunek 6.4: Wahadło podwójne

Klasycznym przykładem zastosowania lagranżowskiego formalizmu jest wahadło podwójne – układ dwóch mas m_1 i m_2 , zawieszonych na niciach o długości l_1 i l_2 , przy czym masa m_1 stanowi punkt zawieszenia wahadła z masą m_2 (por. rysunek 6.4). a dwa równania Lagrange'a Jako współrzędne przyjmujemy dwa kąty: θ_1 i θ_2 , przy których pomocy wyrażamy wszystkie współrzędne kartezjańskie mas. Zagadnienie rozwiązujemy dla małych wychyleń obu wahadeł, pomijając w rozwinięciach funkcji trygonometrycznych wszystkie wyrazy rzędu wyższego niż 2. I tak

⁹Warto wspomnieć, że Lagrange doszedł do nich bez stosowania rachunku wariacyjnego.

$$\begin{aligned}x_1 &= l_1 \sin \theta_1 \approx l_1 \theta_1, \\y_1 &= -l_1 \cos \theta_1 \approx -l_1 \left(1 - \theta_1^2/2\right); \\x_2 &= l_1 \sin \theta_1 + l_2 \sin \theta_2 \approx l_1 \theta_1 + l_2 \theta_2, \\y_2 &= -l_1 \cos \theta_1 - l_2 \cos \theta_2 \approx -l_1 \left(1 - \theta_1^2/2\right) - l_2 \left(1 - \theta_2^2/2\right).\end{aligned}$$

Obliczamy energię kinetyczną

$$T = \frac{1}{2}m_1 (\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m_2 (\dot{x}_2^2 + \dot{y}_2^2) \approx \frac{1}{2}m_1 l_1^2 \dot{\theta}_1^2 + \frac{1}{2}m_2 (l_1 \dot{\theta}_1 + l_2 \dot{\theta}_2)^2$$

i potencjalną

$$V = m_1 g y_1 + m_2 g y_2 \approx -(m_1 l_1 + m_2 l_2)g + (m_1 + m_2)l_1 g \frac{\theta_1^2}{2} + m_2 l_2 g \frac{\theta_2^2}{2}.$$

Lagranżjan układu $L = T - V$ będzie więc równy [przy pominięciu nieistotnej stałej $(m_1 l_1 + m_2 l_2)g$]

$$L = \frac{1}{2}m_1 l_1^2 \dot{\theta}_1^2 + \frac{1}{2}m_2 (l_1 \dot{\theta}_1 + l_2 \dot{\theta}_2)^2 - (m_1 + m_2)l_1 g \frac{\theta_1^2}{2} - m_2 l_2 g \frac{\theta_2^2}{2}, \quad (6.50)$$

a dwa równania Lagrange'a

$$\begin{aligned}\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_1} - \frac{\partial L}{\partial \theta_1} &= 0, \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_2} - \frac{\partial L}{\partial \theta_2} &= 0\end{aligned}$$

prowadzą do układu równań ($M = m_1 + m_2$):

$$\begin{aligned}M l_1 \ddot{\theta}_1 + m_2 l_2 \ddot{\theta}_2 + M g \theta_1 &= 0, \\ l_1 \ddot{\theta}_1 + l_2 \ddot{\theta}_2 + g \theta_2 &= 0.\end{aligned} \quad (6.51)$$

Jest to układ równań jednorodnych, o stałych współczynnikach, który rozwiążemy standardowym podstawieniem

$$\theta_{1,2} = A_{1,2} e^{i\omega t}.$$

Warunek zerowania się wyznacznika głównego układu dostarczy równania na możliwe wartości częstości ω – bliższa analiza wykaże że z czterech możliwych wartości są akceptowalne tylko dwie, a więc liczba częstości własnych układu jest – zgodnie z oczekiwaniami – równa liczbie stopni swobody układu. Po podstawieniu częstości

własnych do układu równań znajdujemy wartości współczynników $A_{1,2}$. Rachunki, choć nietrudne, są dość skomplikowane i dlatego pozostawiamy je dla Dociekliwego i Pilnego Czytelnika.

Równania te znakomicie też działają dla ruchu bez żadnych więzów. Na przykład, dla oscylatora harmonicznego, punktu o masie m , poddanego działaniu siły sprężystej $\mathbf{F} = -k\mathbf{x}$ mamy

$$L = T - V = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$$

i równanie (6.49) ($i = 1$, $q_1 = x$) daje

$$-\frac{d}{dt} \frac{\partial T}{\partial \dot{x}} = -\frac{d}{dt} \frac{dT}{d\dot{x}} = \frac{\partial V}{\partial x} = \frac{dV}{dx},$$

albo

$$-\frac{d}{dt}m\dot{x} = -m\ddot{x} = kx,$$

a więc dobrze nam znane równanie oscylatora harmonicznego z częstością $\omega = \sqrt{k/m}$.

6.4 Kilka zmiennych niezależnych i jedna zmienna zależna

Problem wariacyjny to znalezienie ekstremum funkcjonału, którego postać możemy przyjąć jako

$$\mathcal{J} = \iiint f(v, v_x, v_y, v_z, x, y, z) dx dy dz. \quad (6.52)$$

Funkcja podcałkowa zależy więc od *kilku zmiennych niezależnych*, względem których całkujemy, oraz od jednej (tutaj) zmiennej zależnej $v = v(x, y, z)$. Tak, jak w poprzednich podrozdziałach, wariacja

$$\delta v = \delta v(x, y, z) = V(x, y, z) - v(x, y, z), \quad (6.53)$$

gdzie $v(x, y, z)$ to nasza szukana funkcja, dla której wartość funkcjonału będzie ekstremalna, a $V(x, y, z)$ – funkcja porównawcza, będzie znikać na krańcach przedziałów całkowania. Wariacja funkcjonału, obliczona według schematów identycznych jak w obu poprzednich podrozdziałach, to

$$\delta \mathcal{J} = \iiint \left(\frac{\partial f}{\partial v} \delta v + \frac{\partial f}{\partial v_x} \delta v_x + \frac{\partial f}{\partial v_y} \delta v_y + \frac{\partial f}{\partial v_z} \delta v_z \right) \delta v(x, y, z) dx dy dz = 0. \quad (6.54)$$

Całkowanie przez części, względem każdej z trzech niezależnych zmiennych, i skorzystanie z zerowania się wariacji na krańcach przedziału całkowania prowadzi do

$$\delta \mathcal{J} = \iiint \left(\frac{\partial f}{\partial v} - \frac{d}{dx} \frac{\partial f}{\partial v_x} - \frac{d}{dy} \frac{\partial f}{\partial v_y} - \frac{d}{dz} \frac{\partial f}{\partial v_z} \right) \delta v(x, y, z) dx dy dz = 0 \quad (6.55)$$

a pełna dowolność wariacji $\delta v(x, y, z)$ uprawnia nas do zastąpienia (6.55) równaniem

$$\frac{\partial f}{\partial v} - \frac{d}{dx} \frac{\partial f}{\partial v_x} - \frac{d}{dy} \frac{\partial f}{\partial v_y} - \frac{d}{dz} \frac{\partial f}{\partial v_z} = 0. \quad (6.56)$$

Jako przykład zastosowania tego wariantu równania wariacyjnego możemy rozważyć całkę, względem zmiennej czasowej t , z lagranżjanu drgającej struny. Energia kinetyczna elementu dx struny to $dT = \frac{1}{2} \rho \psi_t^2 dx$; jego energia potencjalna to $dV = \frac{1}{2} \tau \psi_x^2 dx$, gdzie $\psi = \psi(x, t)$ to poprzeczne wychylenie struny; ψ_t oraz ψ_x to odpowiednio pochodna czasowa (prędkość) i przestrzenna (przesunięcie) tego wychylenia; ρ to gęstość liniowa struny, a τ – naprężenie (składowa pozioma, stała w przybliżeniu „małych drgań”) działające na oba końce elementu dx . Lagranżjan drgającej struny o długości a to

$$L = \int_0^a \left(\frac{1}{2} \rho \psi_t^2 - \frac{1}{2} \tau \psi_x^2 \right) dx. \quad (6.57)$$

Całka, względem czasu, z lagranżjanu, po przedziale $t \in [t_1, t_2]$ to

$$\mathcal{J} = \int_{t_1}^{t_2} \int_0^a \left(\frac{1}{2} \rho \psi_t^2 - \frac{1}{2} \tau \psi_x^2 \right) dx dt. \quad (6.58)$$

Mamy więc

$$f = f(v_t, v_x) = \frac{1}{2} \rho \psi_t^2 - \frac{1}{2} \tau \psi_x^2;$$

i konsekwentnie

$$\begin{aligned} \frac{d}{dt} \frac{\partial f}{\partial \psi_t} &= \frac{d}{dt} \frac{1}{2} \rho 2 \psi_t = \rho \psi_{tt} \\ \frac{d}{dx} \frac{\partial f}{\partial \psi_x} &= - \frac{d}{dx} \frac{1}{2} \tau 2 \psi_x = -\tau \psi_{xx} \end{aligned}$$

i podstawienie do równania (6.56) daje natychmiast

$$\frac{\partial^2 \psi}{\partial t^2} - \frac{\tau}{\rho} \frac{\partial^2 \psi}{\partial x^2} = 0, \quad (6.59)$$

albo

$$\frac{\partial^2 \psi}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = 0, \quad (6.60)$$

gdzie $c = \sqrt{\tau/\rho}$ to prędkość rozchodzenia się fali w strunie. Warto zastanowić się nad „znikaniem wariacji” na krańcach przedziału całkowania. Znikanie wariacji $v(x, t) = \psi(x, t)$ w chwilach t_1 i t_2 oznacza, że znamy początkowy i końcowy stan struny (w rozdziale trzecim mieliśmy dyspozycji kształt i rozkład prędkości w chwili początkowej – łatwo zrozumieć, że obie informacje są sobie równoważne). Z kolei, znikanie wariacji ψ na krańcach przedziału zmiennej x (punkty 0 i a) oznacza, że – na przykład – oba końce struny są zamocowane. Dopiero dysponując takimi „dodatkowymi” informacjami, jesteśmy w stanie w pełni rozwiązać równanie falowe (6.60) i znaleźć drgania struny $\psi = \psi(x, t)$.

6.5 Kilka zmiennych niezależnych i kilka zmiennych zależnych

To oczywista kompilacja dwóch poprzednich podrozdziałów. Jeżeli zmienne niezależne oznaczyć przez

$$\mathbf{x} \equiv x_1, \dots, x_i, \dots, x_m, \quad (6.61)$$

a zmienne zależne przez

$$\mathbf{y} \equiv y_1(x_1, \dots, x_m), \dots, y_j(x_1, \dots, x_m), \dots, y_n(x_1, \dots, x_m), \quad (6.62)$$

to po połączeniu równań (6.44) i (6.56) warunek ekstremum dla funkcjonału

$$\mathcal{J} = \int_{x_{11}}^{x_{12}} \dots \int_{x_{m1}}^{x_{m2}} f \left(y_1, \dots, y_n, \left. \frac{\partial y_j}{\partial x_i} \right|_{i=1, \dots, m}^{j=1, \dots, n}, x_1, \dots, x_m \right) dx_1, \dots, dx_m \quad (6.63)$$

będzie równoważny n równaniom

$$\frac{\partial f}{\partial y_j} - \sum_{i=1}^m \frac{d}{dx_i} \left(\frac{\partial f}{\partial [\partial y_j / \partial x_i]} \right) = 0, \quad j = 1, 2, \dots, n. \quad (6.64)$$

6.6 Mnożniki Lagrange’a

6.6.1 Szukanie ekstremum w obecności więzów

W podrozdziale 6.3 wprowadziliśmy pojęcie stopni swobody i *współrzędnych uogólnionych*. Te ostatnie, w liczbie równej liczbie stopni swobody, opisują ruch układu, poddanego *więzom*. Istnienie więzów przejawia się formalnie przez istnienie pewnych „dodatkowych” związków, łączących występujące w problemie zmienne.

Metoda *mnożników Lagrange’a* to w pewnym sensie alternatywne podejście do problemu więzów. Zamiast dobierać współrzędne uogólnione, redukując liczbę zmiennych w naszym problemie, możemy rozpatrywać *wszystkie* zmienne, a istnienie więzów uwzględnić przez dołączenie dodatkowych równań.

Metoda mnożników Lagrange’a znajduje zastosowanie nie tylko w rachunku wariacyjnym, ale w „zwykłych” problemach szukania ekstremum funkcji, jeżeli pomiędzy zmiennymi, poprzez które funkcja ta się wyraża, istnieją dodatkowe związki – więzy. Jeżeli rozważyć problem szukania ekstremum dla funkcji $f = f(x, y, z)$, to warunkiem istnienia ekstremum jest jak wiemy

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz = 0. \quad (6.65)$$

Dla w pełni niezależnych zmiennych x, y, z warunek ten implikuje

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0. \quad (6.66)$$

Przypuśćmy teraz, że jednak pomiędzy zmiennymi x, y, z istnieją pewne związki – na przykład *jeden* związek, dany równaniem

$$\phi(x, y, z) = 0, \quad (6.67)$$

z którego mamy

$$d\phi = \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy + \frac{\partial \phi}{\partial z} dz = 0. \quad (6.68)$$

Różniczki zmiennych nie są już niezależne – jedną z nich (np. dz) możemy wyliczyć, znając pozostałe (dx i dy), a jeżeli tak, to implikacja (6.66) przestaje być prawdziwa. Możemy jednak dodać lewe strony równań (6.65) i (6.68), mnożąc to ostatnie przez dowolną (niewiadomą) stałą λ i rozważać sumę różniczek zupełnych

$$df + \lambda d\phi = \left(\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} \right) dx + \left(\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} \right) dy + \left(\frac{\partial f}{\partial z} + \lambda \frac{\partial \phi}{\partial z} \right) dz = 0, \quad (6.69)$$

w której *mnożnik Lagrange’a*, λ , dobieramy tak, aby wyeliminować „przeszkadzający” nam wyraz, a więc aby zachodziło

$$\frac{\partial f}{\partial z} + \lambda \frac{\partial \phi}{\partial z} = 0.$$

Równanie (6.69) przybiera wówczas postać

$$\left(\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} \right) dx + \left(\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} \right) dy = 0, \quad (6.70)$$

co przy niezależności dx i dy prowadzi do

$$\left. \begin{aligned} \frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} &= 0, \\ \frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} &= 0 \end{aligned} \right\} \quad (6.71)$$

– układu równań, który pozwala wyznaczyć warunki (wartości zmiennych x i y) na ekstremum funkcji $f(x, y, z)$, przy dodatkowych więzach, jakie spełniają zmienne x, y, z [równanie (6.67)].

Jako przykład zastosowania techniki mnożników Lagrange'a rozpatrzmy kwantowo-mechaniczny problem cząstki (masa μ), uwięzionej w walcu o promieniu podstawy a i wysokości h . Ścianki walca są nieprzezroczyste – funkcja falowa znika na poboczniczy i podstawach walca. Szukamy stosunku h/a , który dla stanu o najniższej energii cząstki (stanu podstawowego) minimalizuje tę energię. Dodatkowym warunkiem – więzami – jest stałość objętości walca, $V = \pi a^2 h$.

Równanie Schrödingera, jakie spełnia cząstka, rozwiązujemy we współrzędnych cylindrycznych, stosując metodę separacji zmiennych. Mamy, dla $\psi = \psi(\rho, \phi, z) \equiv R(\rho)F(\phi)Z(z)$

$$-\frac{\hbar^2}{2\mu}\Delta\psi = -\frac{\hbar^2}{2\mu}\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\psi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\psi}{\partial\phi^2} + \frac{\partial^2\psi}{\partial z^2}\right] = E\psi(\rho, \phi, z), \quad (6.72)$$

a po wykonaniu różniczkowania, podzieleniu obu stron przez $(-\hbar^2/2\mu)RFZ$ i wprowadzeniu standardowego oznaczenia $k^2 = 2\mu E/\hbar^2$

$$\frac{1}{R}\frac{1}{\rho}\frac{d}{d\rho}\left(\rho\frac{dR}{d\rho}\right) + \frac{1}{F}\frac{1}{\rho^2}\frac{d^2F}{d\phi^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} = -k^2. \quad (6.73)$$

Zależność od kąta ϕ jest harmoniczna:

$$F(\phi) \equiv F_m(\phi) = a_m \cos m\phi + b_m \sin m\phi, \quad (6.74)$$

skąd

$$\frac{1}{\rho^2}\frac{1}{F_m}\frac{d^2F_m}{d\phi^2} = \frac{-m^2}{\rho^2}; \quad m = 0, 1, \dots \quad (6.75)$$

Podobnie

$$\frac{1}{Z}\frac{d^2Z}{dz^2} = -\kappa^2, \quad (6.76)$$

gdzie z uwagi na warunki

$$\psi(\rho, \phi, z = 0) = \psi(\rho, \phi, z = h) = 0$$

mamy

$$Z(z) \equiv Z_n(z) = \sin\left(\frac{n\pi}{h}z\right); \quad \kappa \equiv \kappa_n = \frac{n\pi}{h}, \quad n = 1, 2, \dots \quad (6.77)$$

Uwzględnienie (6.75) i (6.77) w (6.73) prowadzi do

$$\frac{d^2R_{mn}(\rho)}{d\rho^2} + \frac{1}{\rho}\frac{dR_{mn}(\rho)}{d\rho} + \left[(k^2 - \kappa_n^2) - \frac{m^2}{\rho^2}\right]R_{mn}(\rho) = 0, \quad (6.78)$$

a więc do równania Bessela rzędu m , z rozwiązaniem

$$R_{mn}(\rho) = J_m(\gamma\rho), \quad \gamma^2 \equiv k^2 - \kappa_n^2. \quad (6.79)$$

Warunek zerowania się funkcji falowej na poboczniczy walca daje

$$J_m(\gamma a) = 0, \quad \rightarrow \gamma \equiv \gamma_{mp} = \frac{\alpha_{mp}}{a}, \quad p = 1, 2, \dots, \quad (6.80)$$

gdzie α_{mp} to kolejne zero (o numerze p) funkcji Bessela J_m . Drugie z równań (6.79) daje więc

$$k^2 \equiv k_{mnp}^2 = \frac{2\mu E}{\hbar^2} = \gamma_{mp}^2 + \kappa_n^2 = \frac{\alpha_{mp}^2}{a^2} + \frac{n\pi^2}{h^2}; \quad \begin{array}{l} m = 0, 1, \dots \\ n, p = 1, 2, \dots \end{array} \quad (6.81)$$

Stan o najniższej energii E_0 będzie odpowiadał $m = 0$ (wynika to z równania (6.73) i równania (6.75) – przyczynek „azymutalny” do całkowitej energii jest proporcjonalny do m^2 , a wartości pierwszych zer funkcji Bessela J_m rosną wraz ze wskaźnikiem m), oraz najniższym możliwym wartościami pozostałych „liczb kwantowych” $p = n = 1$. Przekształcając równanie (6.81), otrzymujemy szukany związek

$$E_0 = E(a, h) = \frac{\hbar^2}{2\mu} \left(\frac{\alpha_{01}^2}{a^2} + \frac{\pi^2}{h^2} \right), \quad (6.82)$$

gdzie $\alpha_{01} \approx 2,405$ to pierwsze zero funkcji Bessela J_0 . Dołączając do równania (6.82) warunek $V = \pi a^2 h$, otrzymujemy odpowiednik (6.69) ($f \equiv E$; $\phi \equiv V$; $x \equiv a$; $y \equiv h$)

$$\begin{aligned} \left(\frac{\partial E}{\partial a} + \lambda \frac{\partial V}{\partial a} \right) da + \left(\frac{\partial E}{\partial h} + \lambda \frac{\partial V}{\partial h} \right) dh = \\ \left[\frac{\hbar^2}{2\mu} \left(-\frac{2\alpha_{01}^2}{a^3} \right) + \lambda 2\pi a h \right] da + \left[\frac{\hbar^2}{2\mu} \left(-\frac{2\pi^2}{h^3} \right) + \lambda \pi a^2 \right] dh = 0. \end{aligned} \quad (6.83)$$

W zasadzie dalsza procedura powinna obejmować wyznaczenie stałej λ , dla której jedno z wyrażeń w kwadratowych nawiasach równe jest zeru, a następnie – dla tak określonej stałej λ – przyrównanie do zera drugiego wyrażenia w kwadratowych nawiasach. Zabiegi te możemy uprościć rozważając *a fortiori* układ równań

$$\begin{aligned} \frac{\hbar^2}{2\mu} \frac{2\alpha_{01}^2}{a^3} &= \lambda 2\pi a h, \\ \frac{\hbar^2}{2\mu} \frac{2\pi^2}{h^3} &= \lambda \pi a^2. \end{aligned}$$

Dzielimy oba równania stronami i otrzymujemy odpowiedź na nasze pytanie

$$\frac{h}{a} = \frac{\sqrt{2\pi}}{\alpha_{01}} \approx 1,85.$$

Taki właśnie stosunek wysokości do promienia podstawy walca (przy ustalonej¹⁰ objętości $V = \pi a^2 h$) zapewnia najniższą wartość energii stanu podstawowego uwiecznionej w nim cząstki.

Jak wynika z prezentowanego przykładu, udało się go rozwiązać *bez* wyznaczania wartości liczbowej mnożnika Lagrange'a λ . Często tak bywa i dlatego też metoda jest nazywana *metodą nieoznaczonych (nieokreślonych) mnożników Lagrange'a*. Nie znaczy to jednak, że mnożniki nie zawierają w sobie ilościowej (i ciekawej) informacji. Zobaczmy to kolejnym podrozdziałem, gdzie mnożniki Lagrange'a współuczestniczą w określeniu sił reakcji.

6.6.2 Mnożniki Lagrange'a w kontekście wariacyjnym

Przypuśćmy, że rozważamy przypadek jednej zmiennej niezależnej, którą jest czas t i m zmiennych zależnych, $x_1(t), \dots, x_i(t), \dots, x_m(t) \equiv \mathbf{x}$, opisujących ruch naszego układu. Pomiedzy współrzędnymi \mathbf{x} istnieje r związków; ich równania zapisujemy jako

$$\phi_k[x_1(t), \dots, x_i(t), \dots, x_m(t), t] = 0, \quad l = 1, \dots, r. \quad (6.84)$$

Pozostając w kontekście „mechanicznym”, rozważmy problem minimalizacji całki z lagranżajnu [por. (6.46)], a więc warunek zerowej wariacji z wyrażenia

$$\delta L = \delta \int_{t_1}^{t_2} L dt = \delta \int_{t_1}^{t_2} L(\mathbf{x}, \dot{\mathbf{x}}, t) dt = 0. \quad (6.85)$$

Nie jest to *de facto* „prawdziwy” lagranżjan, bo jego argumenty, zmienne \mathbf{x} , są powiązane r związkami (6.84). Dlatego też, wariacje poszczególnych współrzędnych nie są niezależne i zerowanie się wariacji z funkcjonału (6.85) nie implikuje równań Eulera-Lagrange'a. Możemy jednak każde z równań (6.84) przemnożyć przez pewien (nieokreślony) czynnik λ_k , dodać do siebie tak utworzone wyraży, a uzyskaną sumę poddać operacji całkowania. Otrzymamy wyrażenie

$$\int_{t_1}^{t_2} \sum_{k=1}^r \lambda_k \phi_k(\mathbf{x}, t) dt. \quad (6.86)$$

W odróżnieniu od dyskutowanej w poprzednim podrozdziale metody mnożników, w przypadku wariacji z więzami okazuje się, że mnożniki λ_k mogą zależeć od zmiennej niezależnej t – $\lambda_k = \lambda_k(t)$. Z kolei, zależność równań więzów od czasu występuje tylko w przypadku tzw. *więzów reonomicznych*, z którymi raczej rzadko miewamy do czynienia w praktyce (większość problemów mechanicznych dotyczy tzw. *więzów skleronomicznych*¹¹, których równania nie zależą *explicite* od czasu).

¹⁰Zauważmy, że bez tego dodatkowego warunku (stałość objętości walca) pytanie: jakie proporcje powinien mieć walec, aby zapewnić minimalną energię stanu podstawowego, nie miałyby sensu.

¹¹Z greckiego: *rheó, rhein* – płynąć; *skléros* – twardy.

Wariacja wyrażenia (6.86) jest też równa zero

$$\delta \int_{t_1}^{t_2} \sum_{k=1}^r \lambda_k \phi_k(\mathbf{x}, t) dt = 0,$$

a więc możemy rozpatrywać, tak jak w podrozdziale poprzednim, problem wariacyjny „rozszerzony” o równania więzów

$$\delta L = \delta \int_{t_1}^{t_2} L dt \rightarrow \delta \int_{t_1}^{t_2} \left[L(\mathbf{x}, \dot{\mathbf{x}}, t) + \sum_{k=1}^r \lambda_k \phi_k(\mathbf{x}, t) \right] dt = 0, \quad (6.87)$$

w którym mnożniki λ_k dobieramy tak, aby wyeliminować zależne od siebie wariacje współrzędnych. Procedury analogiczne do tych, które stosowaliśmy w podrozdziale 6.3 prowadzą do układu m równań (dla każdej zmiennej zależnej x_i mamy jedno równanie)

$$\frac{\partial}{\partial x_i} \left(L - \sum_{k=1}^r \lambda_k \phi_k \right) - \frac{d}{dt} \frac{\partial}{\partial \dot{x}_i} \left(L - \sum_{k=1}^r \lambda_k \phi_k \right) = 0, \quad i = 1, \dots, m \quad (6.88)$$

Pozostawiamy Czytelnikowi wykazanie, że w przypadku kiedy w równaniach więzów nie występują pochodne czasowe współrzędnych, tzn. kiedy

$$\frac{\partial \phi_k}{\partial \dot{x}_i} = 0, \quad \begin{array}{l} k = 0, \dots, r \\ i = 1, \dots, m \end{array}$$

zmodyfikowane równania Lagrange’a [por. (6.48)] przybierają postać

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = \sum_{k=1}^r \lambda_k \frac{\partial \phi_k}{\partial x_i}. \quad (6.89)$$

Zastanówmy się przez chwilę nad fizycznym sensem wyrażenia po prawej stronie równania (6.89). Załóżmy, że posługujemy się układem współrzędnych kartezjańskich, w którym wszystkie x_i oznaczają odległości (długości). Występujący po lewej stronie wyraz $-\partial L/\partial x_i$ jest równy – dla energii kinetycznej T zależnej tylko od pochodnych względem czasu współrzędnych x_i – pochodnej energii potencjalnej względem x_i (ze znakiem ujemnym) $-\partial V/\partial x_i$ a więc reprezentuje siłę „zewnętrzna” (jej i -tą składową). Analogicznie, wyrażenia $\lambda_k \partial \phi_k/\partial x_i$ będą reprezentować *siły reakcji* – a konkretnie i -tą składową reakcji, pochodzącej od więzów zadanych równaniem $\phi_k(\mathbf{x}, t) = 0$.

Prostym przykładem zastosowania techniki mnożników Lagrange’a może być problem ruchu wahadła matematycznego, punktu o masie m , zawieszono na nieważkiej nici o stałej długości równej l . To ostatnie sformułowanie to nasze równanie więzów

$$\phi_1(r) = r - l = 0, \quad (6.90)$$

gdzie r jest odległością masy od punktu zawieszenia. Jako drugą współrzędną wybieramy kąt θ (układ współrzędnych biegunowych). Energia kinetyczna punktu to

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m(v_r^2 + v_\theta^2) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2);$$

energia potencjalna z kolei to

$$V = -mgr \cos \theta.$$

Wyrażenie podcałkowe (6.87) ma więc postać

$$\begin{aligned} L(r, \theta, \dot{r}, \dot{\theta}) + \lambda_1 \phi_1(r) &= T - V + \lambda_1 \phi_1(r) = \\ &= \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + mgr \cos \theta + \lambda_1 \phi_1(r). \end{aligned} \quad (6.91)$$

Równania Lagrange'a (tzw. równania Lagrange'a *pierwszego rodzaju*) – odpowiednik (6.89) – to

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = \lambda_1 \frac{\partial \phi_1}{\partial r} = \lambda_1, \quad (6.92)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = \lambda_1 \frac{\partial \phi_1}{\partial \theta} = 0, \quad (6.93)$$

a po podstawieniu z (6.91) za L i wykonaniu różniczkowania dostajemy dwa równania

$$m \frac{d\dot{r}}{dt} - mr\dot{\theta}^2 - mg \cos \theta = \lambda_1, \quad (6.94)$$

$$m \frac{d(r^2\dot{\theta})}{dt} + mgr \sin \theta = 0. \quad (6.95)$$

W otrzymanych równaniach odległość r od punktu zawieszenia jest jednak ustalona przez równanie więzów (6.90) – mamy $r = l$ i oczywiście $\dot{r} = 0$. Ostatecznie więc nasze dwa równania przybierają postać

$$ml\dot{\theta}^2 + mg \cos \theta = -\lambda_1, \quad (6.96)$$

$$ml^2\ddot{\theta} + mgl \sin \theta = 0. \quad (6.97)$$

Drugie równanie – (6.97) – to właściwe równanie ruchu wahadła, określające wychylenie (kąt θ) w funkcji czasu. Dla małych wartości θ równanie to możemy efektywnie rozwiązać, w przybliżeniu ruchu harmonicznego $\sin \theta \approx \theta$. Natomiast pierwsze równanie – (6.96) – to równanie w którym określona jest siła reakcji więzów (nici) λ_1 . Siła ta, skierowana wzdłuż nici, równoważy odpowiednią składową siły ciężkości i „dostarcza” poruszającej się masie odpowiedniego przyspieszenia dośrodkowego.

Z omówionego przykładu wynika, że stosowanie „prostszych” równań Lagrange’a – tzn. równań drugiego rodzaju, w których występują już tylko współrzędne uogólnione, w liczbie równej liczbie stopni swobody układu, będzie wskazane, jeżeli nie interesują nas siły reakcji więzów. W przeciwnym wypadku, jesteśmy zmuszeni do użycia równań Lagrange’a pierwszego rodzaju, w których zawarte są równania drugiego rodzaju, a także ilościowy opis sił reakcji więzów, którym poddany jest układ. W zastosowaniach praktycznych („inżynierskich”) informacje o siłach reakcji będą ważniejsze od samego opisu ruchu.

6.7 Podsumowanie

Ten skromny rozdział o rachunku wariacyjnym zawiera podstawowe wiadomości, związane z bardzo podstawowym i bardzo „starym” problemem fizyki: jak zminimalizować pewne wielkości, które nie są prostą funkcją pewnej (ciągłej) zmiennej, ale raczej wyrażają się za pomocą często dość skomplikowanej formuły, która okazuje się całką. Całką, której wartość dodatkowo zależy od wyboru podpowierzchni (często po prostu krzywej), na której jest ona obliczana. Wartość całki zależy też od tej *zmiennej niezależnej* (bądź: *zmiennych niezależnych*, ale ciężar zależności leży na funkcji (i jej pochodnej, bądź pochodnych) tej zmiennej niezależnej, określającej wybór drogi całkowania.

Rozwiązanie problemu wariacyjnego przez Eulera i Lagrange’a umożliwiło całkowicie „nowe” sformułowanie mechaniki klasycznej w kontekście teoretycznym. W rozdziale podajemy podstawowe równania tej mechaniki lagranżowskiej, która w latach 30. XIX w. została uzupełniona o bardzo istotny rozdział – równania i formalizm Hamiltona. Ten ostatni jest wyprowadzony z zasady wariacyjnej. Głębsze wniknięcie w strukturę tego formalizmu (bardzo polecane) doprowadzi Dociekliwego Czytelnika jeszcze raz do ...operatorów samosprzężonych. Można bowiem wykazać, że każdy problem sformułowany według *zasady wariacyjnej* prowadzi do operatorów, które mają właśnie własność samosprzężoności, a o których mówiliśmy w rozdziale trzecim. Formalizm wariacyjny (hamiltonowski) powstał na gruncie fizyki klasycznej, ale – nie będzie to nas dziwić – odniósł jeszcze chyba większe sukcesy w sformułowaniu mechaniki kwantowej.

Autor nie odmówił sobie przyjemności przedyskutowania, dość szczegółowego, słynnego problemu brachistochrony, który z końcem XVII w. zaprzętał umysły największych matematyków i fizyków. Jego rozwiązanie można wprawdzie uzyskać (i uzyskano), nie posługując się właśnie rachunkiem wariacyjnym, ale niewątpliwie zastosowanie tego formalizmu jest w tym przypadku bardzo efektowne.

Kolejne podrozdziały to przygotowanie Czytelnika do kursu podstaw fizyki teoretycznej, który powinien rozwinąć treści tych podrozdziałów, potraktowane tutaj skrótowo.

Literatura

- [1] Golfand, I. M., Fomin, S. V., *Calculus of Variations*, Dover Publ., 1991
- [2] Byron, F. W., Fuller, R. W.: *Mathematics of Classical and Quantum Physics*, Dover Publ. 1992 (anglojęzyczna reedycja cytowanego wcześniej polskiego tłumaczenia)
- [3] Arfken, G.: *Mathematical Methods for Physicists, 4th Ed.* New York and London, Academic Press 1995