

Andrzej J. Maciejewski

Wykłady z mechaniki

Zielona Góra, 16 grudnia 2013

1. Wstęp

Niniejsze tekst powstał w oparciu o wykłady jakie wygłosiłem w Instytucie Matematyki i Informatyki UMK w Toruniu w listopadzie 2013 r. Pierwszą jego wersję, w oparciu o notatki z wykładu, opracowała Pani mgr. Marta Kowalczyk za co jej serdecznie dziękuję.

W pierwszym odruchu planowałem wygłosić cykl wykładów przedstawiających wybrane metody badania całkowalności i niecałkowalności jako, że tymi zagadnieniami aktualnie się zajmuje. Jednak później zdecydowałem się ograniczyć do bardzo klasycznego materiału. O ile formalny język mechaniki (tzn. geometria różniczkowa) jest łatwo „trawiony” przez słuchaczy z dobrym przygotowaniem matematycznych, to zdecydowanie trudniej jest przekazać wiedzę fizyczną, która dla wykładowcy jest wiedzą „oczywistą”, ale dla słuchaczy już niekoniecznie.

2. Podstawy

Celem wykładu jest przedstawienie podstawowych założeń i pojęć mechaniki Newtona. Cała mechanika klasyczna opiera się na mechanice Newtona. Z jednej strony, formalizmy Lagrange’a i Hamiltona, opisują tylko pewne klasy układów, które opisuje mechanika Newtona, z drugiej zaś, są one opisem pięknej geometrycznej struktury mechaniki, która nie jest widoczna w opisie Newtona.

§ 1. Czasoprzestrzeń Newtona. Przestrzeń w której odbywa się ruch jest trójwymiarową rzeczywistą przestrzenią euklidesową \mathbb{E}^3 . Punkty tej przestrzeni będziemy oznaczać dużymi łacińskimi literami: $P, Q, R \dots$. Wyodróżnijmy jeden punkt $O \in \mathbb{E}^3$. Wówczas dowolny punkt $P \in \mathbb{E}^3$ jest jednoznacznie określony przez swój wektor wodzący $\mathbf{r}(P) = \vec{OP} = P - O$. Z definicji, wektor $\mathbf{r}(P)$ jest elementem rzeczywistej przestrzeni unitarnej stowarzyszonej z \mathbb{E}^3 . Będziemy ją oznaczać \mathbb{R}^3 , pamiętając, że jest ona wyposażona w iloczyn skalarny. Dla wektorów $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ ich iloczyn skalarny oznaczmy standardowo $\mathbf{x} \cdot \mathbf{y}$. Dalej $\|\mathbf{x}\| := \sqrt{\mathbf{x} \cdot \mathbf{x}}$ oznaczać będzie długość wektora \mathbf{x} .

Czas jest jednowymiarowy, utożsamiamy go z prostą rzeczywistą \mathbb{R} . Czasoprzestrzeń Newtona, to $\mathbb{E}^3 \times \mathbb{R}$, rozpatrywana jako przestrzeń afiniczna.

Jeżeli punkt $P \in \mathbb{E}^3$ się porusza, to jego wektor wodzący jest funkcją czasu $\mathbf{r} = \mathbf{r}(t)$. Prędkość punktu to $\mathbf{v} := \dot{\mathbf{r}}$, a jego przyspieszenie dane jest przez $\mathbf{a} = \ddot{\mathbf{r}} = \dot{\mathbf{v}}$.

Punkt materialny, to punkt w \mathbb{E}^3 , który posiada masę $m > 0$. Dynamiczny stan punktu materialnego P opisywany jest parą $(P, \mathbf{v}) \in \mathbb{E}^3 \times \mathbb{R}^3$. Ponieważ zawsze będziemy posługiwać się jakimś układem, to stan będziemy utożsamiać z parą $(\mathbf{r}, \mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3$. Zbiór wszystkich stanów nazywany

jest przestrzenią stanów lub przestrzenią fazową. Dla punktu materialnego przestrzenią stanów jest $\mathbb{E}^3 \times \mathbb{R}^3$.

Jeżeli rozpatrujemy układ składający się z n punktów materialnych P_i o masach $m_i, i = 1, \dots, n$, to wprowadzamy następujące oznaczenia

$$\mathbf{r} := (\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{E}^{3n}, \quad \mathbf{v} := (\mathbf{v}_1, \dots, \mathbf{v}_n) \in \mathbb{R}^{3n},$$

gdzie $\mathbf{r}_i := \overset{\#}{\check{O}}P_i$ oraz $\mathbf{v}_i := \dot{\mathbf{r}}_i$ dla $i = 1, \dots, n$. Wówczas stanem układu jest para $(\mathbf{r}, \mathbf{v}) \in \mathbb{E}^{3n} \times \mathbb{R}^{3n}$. Przestrzeń

$$\mathbb{E}^{3n} = \underbrace{\mathbb{E}^3 \times \dots \times \mathbb{E}^3}_{n \text{ razy}},$$

ze współrzędnymi $\mathbf{r} := (\mathbf{r}_1, \dots, \mathbf{r}_n)$ nazywana jest przestrzenią konfiguracyjną układu. Jeżeli chcemy wykluczyć zderzenia, to musimy zmniejszyć przestrzeń konfiguracyjną o następujący zbiór:

$$\mathcal{C} := \bigcup_{1 \leq i < j \leq n} \{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{E}^{3n} \mid \mathbf{r}_i = \mathbf{r}_j\}. \quad (2.1)$$

§2. Wielkości charakteryzujące układy punktów materialnych. Prócz stanu (\mathbf{r}, \mathbf{v}) punktu materialnego o masie m wprowadza się następujące wielkości:

- pęd punktu $\mathbf{p} := m\mathbf{v}$,
- moment pędu $\mathbf{c} := \mathbf{r} \times \mathbf{p}$,
- energia kinetyczna

$$T = \frac{1}{2}m\mathbf{v} \cdot \mathbf{v} = \frac{1}{2m}\mathbf{p} \cdot \mathbf{p},$$

Dla układu n punktów o masach m_i i stanach $(\mathbf{r}_i, \mathbf{v}_i)$ wprowadza się wielkości która charakteryzują układ jako całość. Są to

- masa układu:

$$m := \sum_{i=1}^n m_i,$$

- środek masy układu:

$$\mathbf{R} = \frac{1}{m} \sum_{i=1}^n m_i \mathbf{r}_i,$$

- całkowity pęd układu:

$$\mathbf{p} := \sum_{i=1}^n \mathbf{p}_i = \sum_{i=1}^n m_i \mathbf{v}_i,$$

- całkowity moment pędu układu:

$$\mathbf{c} := \sum_{i=1}^n \mathbf{c}_i = \sum_{i=1}^n \mathbf{r}_i \times \mathbf{p}_i,$$

- energia kinetyczna układu :

$$T = \frac{1}{2} \sum_{i=1}^n m_i \mathbf{v}_i \cdot \mathbf{v}_i = \frac{1}{2} \sum_{i=1}^n \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{m_i}.$$

3. Prawa Newtona

Układ odniesienia w mechanice punktu jest to układ współrzędnych afinicznych w czasoprzestrzeni Newtona $\mathbb{E}^3 \times \mathbb{R}$. Układ taki określony jest przez swój początek $(O, t_0) \in \mathbb{E}^3 \times \mathbb{R}$ oraz wybór bazy $\{e_1, e_2, e_3, e_4\}$ w $\mathbb{R}^3 \times \mathbb{R}$. Zazwyczaj określamy go podając parę $\mathcal{J} = (O, \{e_1, e_2, e_3\})$. Oznacza to, że przyjmujemy $e_4 = (0, 0, 0, 1)$.

Niech $\mathcal{F} = (P, \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\})$ będzie innym układem odniesienia, a $Q \in \mathbb{E}^3$ dowolnym punktem. Dla tego punktu mamy określone dwa wektory wodzące $\mathbf{q} := \overset{\#}{\mathcal{O}}Q$ oraz $\mathbf{r} := \overset{\#}{\mathcal{P}}Q$. Oczywiście zachodzi

$$\mathbf{q} = \mathbf{x} + \mathbf{r}, \quad \text{gdzie} \quad \mathbf{x} := \overset{\#}{\mathcal{O}}P, \quad (3.1)$$

zobacz rys. Jeżeli punkt Q oraz cały układ \mathcal{F} się porusza, to zależność powyższa zależy istotnie od tego jakimi funkcjami czasu są wektory $\mathbf{b}_1, \mathbf{b}_2$ i \mathbf{b}_3 . Jest ona szczególnie prosta, gdy początki układów poruszają się względem siebie ze stałą prędkością, tzn. $\ddot{\mathbf{x}} = \mathbf{0}$, oraz gdy orientacja układów nie zależy od czasu, tzn. gdy elementy macierzy $[A_{ij}]$ określającej przejście z bazy $\{e_1, e_2, e_3\}$ do bazy $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$

$$\mathbf{b}_i = \sum_{j=1}^3 A_{j,i} \mathbf{e}_j, \quad \text{dla} \quad 1 \leq i \leq 3,$$

nie zależą od czasu. Mówimy wtedy, że układy odniesienia \mathcal{J} i \mathcal{F} poruszają się względem siebie ruchem jednostajnym prostoliniowym.

W mechanice Newtona podstawową rolę odgrywają dwa pojęcia. Pierwsze z nich to *inercjalny układ odniesienia*. Jest to pewien układ odniesienia $\mathcal{J} = (O, \{e_1, e_2, e_3\})$, którego istnienie się postuluje. Każdy inny układ odniesienia $\mathcal{F} = (P, \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\})$, który porusza się względem \mathcal{J} ze stałą prędkością jest również układem inercjalnym.

Fizyczna motywacja wyróżnienia inercjalnych układów odniesienia jest następująca. Chcemy opisać ruch punktu lub punktów w sytuacji gdy żaden czynnik zewnętrzny na nie nie działa. To czy dany układ jest inercjalny czy też nie zależy od zjawiska które badamy.

§ 3. Drugie prawo Newtona. Prawo to można formułować w różny sposób.

Prawo 1 (Newton). Jeżeli na punkt materialny działa siła \mathbf{F} , to jego ruch opisywany jest równaniem

$$m\ddot{\mathbf{r}} = \mathbf{F}. \quad (3.2)$$

Często równanie (3.2), dla punktów materialnych o stałej masie, zapisuje się w postaci

$$\dot{\mathbf{p}} = \mathbf{F}. \quad (3.3)$$

W zasadzie to prawo można uznać za definicję siły. W praktyce siła znana jest z doświadczenia.

Gdy na punkt materialny działa siła \mathbf{F} , to wielkość $\mathbf{M} := \mathbf{r} \times \mathbf{F}$ nazywamy momentem tej siły.

Jeżeli rozpatrujemy układ n punktów materialnych, to zamiast jednego równania Newtona mamy układ równań

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i, \quad i = 1, \dots, n. \quad (3.4)$$

Wypadkowa siła działająca na układ to

$$\mathbf{F} = \sum_{i=1}^n \mathbf{F}_i.$$

Układ nazywamy izolowanym, gdy wypadkowa siła działająca na niego znika, tzn. gdy $\mathbf{F} = \mathbf{0}$.

§ 4. Pierwsze prawo Newtona. Można je sformułować następująco.

Prawo 2 (Newton). *Jeżeli na punkt materialny nie działa żadna siła lub siły działające równoważą się, to w dowolnym układzie inercyjnym punkt ten pozostaje w spoczynku lub porusza się ruchem jednostajnym prostoliniowym.*

Podstawowym pytaniem o ciało jest to o ruch, dynamikę ciała. Pierwsze prawo Newtona odpowiada na pytanie jak porusza się ciało, na które nie działa żadna siła lub siły działające na nie się równoważą. Problem stanowi kwestia względem jakiego układu odniesienia opisujemy ruch tego ciała. Rzeczywiście zakładając, że całą dynamikę ciała w istotny sposób nie spowodowany żadną siłą, w pewnym układzie odniesienia K_0 , będziemy opisywać równaniem $\ddot{\mathbf{r}} = \mathbf{0}$, to pierwsze prawo Newtona będzie nieprawdziwe w pewnym układzie odniesienia K , który będzie się poruszał ruchem przyspieszonym względem K_0 . Dlatego wyróżniamy pewne specjalne układy odniesienia, dla których ruch ciała nie spowodowany w istotny sposób żadną siłą będzie zawsze jednostajny (w szczególności ruch ze stałą prędkością odbywa się po prostej). Pierwsze prawo Newtona definiuje bardzo ważną klasę układów- układów inercyjnych. Istnienie układu inercyjnego jest postulatem. Zauważmy, że nie istnieje jakiś jeden szczególny „układ odniesienia inercyjny” oraz jeżeli istnieje jeden układ inercyjny, to jest ich nieskończenie wiele.

Lemat 1. *Niech K będzie inercyjnym układem odniesienia. Wówczas dowolny inny układ K' , który porusza się względem K ze stałą prędkością \mathbf{w} jest inercyjny.*

Dowód. Załóżmy, że punkt materialny opisujemy wektorami wodzącymi $\mathbf{r}(t)$ oraz $\mathbf{r}'(t)$ odpowiednio w układach K i K' . Wówczas równanie $\mathbf{r}'(t) = \mathbf{r}(t) - \mathbf{w}t$ opisuje ruch punktu materialnego względem układu odniesienia K' . Dwukrotnie różniczkując obustronnie powyższe równanie otrzymujemy, że $\ddot{\mathbf{r}}' = \ddot{\mathbf{r}} = \mathbf{0}$, co dowodzi tezy. \square

Definicja 1. *Układ odniesienia poruszający się względem inercyjnego układu odniesienia ruchem przyspieszonym nazywamy nieinercyjnym układem odniesienia.*

§ 5. Trzecie prawo Newtona. Prawo to brzmi następująco.

Prawo 3. *Jeżeli jedno ciało działa na drugie pewną siłą, to drugie działa na pierwsze siłą o takiej samej wartości bezwzględnej, takim samym kierunku i przeciwnym zwrocie.*

Trzecie prawo Newtona stanowi o dwóch ciałach, parę sił „akcji i reakcji” w sensie tego prawa stanowią siły przyłożone do dwóch różnych ciał a każda z sił stanowiąca tę parę działa na jedno ciało. O ruchu, każdego z tych punktów materialnych, o których mówi trzecie prawo Newtona, stanowi wypadkowa wszystkich sił na nie działających (np. skoczek odbija się od belki a belka od skoczka, ale tylko skoczek leci, a belka pozostaje nieruchoma).

§ 6. Prawo powszechnego ciążenia Newtona. Prawo fizyczne bazujące na obserwacjach empirycznych Newtona mówi, że:

Prawo 4. *Każde dwa punkty materialne przyciągają się z siłami, których wielkość jest wprost proporcjonalna do iloczynu ich mas a odwrotnie proporcjonalna do kwadratu odległości między nimi, tj.*

$$\mathbf{F}_{1,2} = G \frac{m_1 m_2}{r_{12}^3} (\mathbf{r}_1 - \mathbf{r}_2), \quad (3.5)$$

$$\mathbf{F}_{2,1} = -\mathbf{F}_{1,2}, \quad (3.6)$$

gdzie G jest stałą grawitacji oraz $r_{12} = \|\mathbf{r}_1 - \mathbf{r}_2\|$.

§ 7. Zasada przyczynowości Newtona-Laplace’a. Zasada ta mówi, że stan dynamiczny punktu materialnego $(\mathbf{r}_0, \mathbf{v}_0)$ w dowolnej chwili czasu t_0 określa jednoznacznie jego ruch dla $t > t_0$ i dla $t < t_0$. Wynika ona z eksperymentów.

Można ją również uzasadnić wykorzystując twierdzenie o istnieniu i jednoznaczności dla równań różniczkowych zwyczajnych¹.

Twierdzenie 1 (Istnienie i jednoznaczność). *Niech $I \subset \mathbb{R}$ oraz $\Omega \subset \mathbb{R}^n$ będą zbiorami otwartymi, a $\mathbf{f} \in C^1(I \times \Omega, \mathbb{R}^n)$. Wówczas dla każdego $(t_0, \mathbf{x}_0) \in I \times \Omega$, zagadnienie początkowe*

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t)) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases} . \quad (3.7)$$

ma dokładnie jedno nierozszerzalne rozwiązanie $\mathbf{u}(\cdot; t_0, \mathbf{x}_0) : I_{(t_0, \mathbf{x}_0)} \rightarrow \Omega$.

Maksymalny odcinek istnienia rozwiązania $I_{(t_0, \mathbf{x}_0)} = (t^-(t_0, \mathbf{x}_0), t^+(t_0, \mathbf{x}_0))$ jest otwarty. Ponadto zachodzi alternatywa:

1. albo $t^-(t_0, \mathbf{x}_0) = \inf I$ i $t^+(t_0, \mathbf{x}_0) = \sup I$;

¹ Zobacz: Herbert Amann, Ordinary Differential Equations: An Introduction to Non-linear Analysis, Walter de Gruyter, Berlin-New-York 1990.

2. albo

$$\lim_{t \rightarrow t^\pm(t_0, \mathbf{x}_0)} \min\{\text{dist}(u(t; t_0, \mathbf{x}_0), \partial\Omega), \|u(t; t_0, \mathbf{x}_0)\|^{-1}\} = 0,$$

gdzie dodatkowo przyjmujemy $\text{dist}(\mathbf{x}, \emptyset) = +\infty$.

Założmy, że dla pewnej chwili czasu t_0 znamy stan punktu materialnego o masie m . Innymi słowy, znamy położenie $\mathbf{r}(t_0) = \mathbf{r}_0$ oraz prędkość $\mathbf{v}(t_0) = \mathbf{v}_0$ tego punktu. Dodatkowo założmy, że znamy siłę $\mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t)$ działającą na ten punkt materialny. Chcemy znaleźć położenie $\mathbf{r}(t)$ i prędkość $\mathbf{v}(t)$ punktu w chwili $t \neq t_0$, tj. wyznaczyć stan układu w chwili $t \neq t_0$. Z drugiego prawa Newtona wynika, że $m\ddot{\mathbf{r}} = \mathbf{F} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t)$. Zatem problem sprowadza się do rozwiązywania zagadnienia początkowego postaci

$$\begin{cases} m\ddot{\mathbf{r}}(t) &= \mathbf{F}(\mathbf{r}(t), \dot{\mathbf{r}}(t), t) \\ \mathbf{r}(t_0) &= \mathbf{r}_0 \\ \dot{\mathbf{r}}(t_0) &= \dot{\mathbf{r}}_0 \end{cases}.$$

Na podstawie twierdzenia 1 otrzymujemy, że istnieje dokładnie jedno rozwiązanie $\mathbf{r} : I_{(t_0, \mathbf{r}_0, \dot{\mathbf{r}}_0)} \rightarrow \Omega$ powyższego zagadnienia początkowego, zatem znamy stan układu dla $t \in I_{(t_0, \mathbf{r}_0, \dot{\mathbf{r}}_0)}$. Dlatego znajomość sił działających na punkt materialny oraz stanu układu w pewnej chwili jednoznacznie wyznacza stan układu w otoczeniu tej chwili.

4. Dynamika układu punktów materialnych

Rozważmy układ n punktów materialnych $\mathbf{r}_1, \dots, \mathbf{r}_n$ o masach odpowiednio m_1, \dots, m_n . Niech masa wszystkich punktów wynosi m . Zakładamy, że punkty oddziałują ze sobą binarnie. To znaczy, siła z jaką punkt k -ty działa na i -ty wynosi $\mathbf{F}_{i,k} = -\mathbf{F}_{k,i}$. Zatem wypadkowa sił wewnętrznych działająca na i -ty punkt wynosi

$$\mathbf{F}_i = \sum_{\substack{1 \leq k \leq n \\ k \neq i}} \mathbf{F}_{i,k}. \quad (4.1)$$

Wniosek 1. Wypadkowa wszystkich sił wewnętrznych znika, tzn. zach

$$\mathbf{F} := \sum_{i=1}^n \mathbf{F}_i = \mathbf{0}.$$

Dowód. Połóżmy $\mathbf{F}_{i,i} := \mathbf{0}$ dla $1 \leq i \leq n$. Wówczas

$$\mathbf{F}_i = \sum_{k=1}^n \mathbf{F}_{i,k} \quad \text{dla } 1 \leq i \leq n,$$

oraz

$$\mathbf{F} = \sum_{i,k=1}^n \mathbf{F}_{i,k} = - \sum_{i,k=1}^n \mathbf{F}_{k,i}.$$

Zatem $\mathbf{F} = -\mathbf{F}$, co kończy dowód. \square

Założmy dodatkowo, że na i punkt działa siła zewnętrzna \mathbf{K}_i . Rozpatrujemy nasz układ w pewnym układzie inercyjnym. Przez

$$\mathbf{K} := \sum_{i=1}^n \mathbf{K}_i,$$

będziemy oznaczać wypadkową sił zewnętrznych działających na nasz układ.

Mówimy, że nasz układ jest układem izolowanym (zamkniętym) gdy $\mathbf{K}_i = \mathbf{0}$, dla $1 \leq i \leq n$.

Równania Newtona n punktów materialnych mają następującą postać:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i + \mathbf{K}_i, \quad i = 1, \dots, n. \quad (4.2)$$

Poniższy lemat mówi, że środek masy \mathbf{R} naszego układu porusza się tak, jak punkt materialny o masie m pod wpływem siły \mathbf{K} .

Lemat 2 (O ruchu środka masy). *Środek masy \mathbf{R} układu n punktów materialnych spełnia równanie $m\ddot{\mathbf{R}} = \mathbf{K}$. W szczególności, środek masy izolowanego układu porusza się ruchem jednostajnym prostoliniowym.*

Dowód. Rzeczywiście, mamy

$$m\ddot{\mathbf{R}} = \sum_{i=1}^n m_i \ddot{\mathbf{r}}_i = \sum_{i=1}^n (\mathbf{F}_i + \mathbf{K}_i) = \mathbf{F} + \mathbf{K} = \mathbf{K},$$

w ostatniej równości skorzystaliśmy z Wniosku 1. □

Na podstawie lematu 2 możemy wprowadzić pęd środka masy \mathbf{P} . Jest pęd punkt materialny o masie m , którego wektorem wodzącym jest \mathbf{R} . Zatem

$$\mathbf{P} = m\dot{\mathbf{R}} = \frac{d}{dt} m\mathbf{R} = \frac{d}{dt} \sum_{i=1}^n m_i \mathbf{r}_i = \sum_{i=1}^n m_i \dot{\mathbf{r}}_i = \mathbf{p}.$$

Inaczej mówiąc, pęd środka masy jest równy całkowitemu pędowi układu. Z lematu 2 wynika, że $\dot{\mathbf{P}} = \mathbf{K}$.

Wniosek 2 (Zasada zachowania pędu). *Jeżeli $\mathbf{K} = \mathbf{0}$, to całkowity pęd układu jest stały.*

Wniosek 3. *Jeżeli $\mathbf{K} = \mathbf{0}$, to ruch środka masy opisany jest równaniem*

$$\mathbf{R}(t) = t \frac{\mathbf{P}}{m} + \mathbf{R}(0).$$

Lemat 3. *Dla $1 \leq i \leq n$, zachodzi*

$$\dot{\mathbf{c}}_i = \mathbf{M}_i = \mathbf{r}_i \times (\mathbf{F}_i + \mathbf{K}_i),$$

oraz

$$\dot{\mathbf{c}} = \mathbf{M} := \sum_{i=1}^n \mathbf{M}_i.$$

Dowód. Ponieważ $\dot{\mathbf{r}}_i \times \mathbf{p}_i = \mathbf{0}$, to $\dot{\mathbf{c}}_i = \dot{\mathbf{r}}_i \times \mathbf{p}_i + \mathbf{r}_i \times \dot{\mathbf{p}}_i = \mathbf{r}_i \times (\mathbf{F}_i + \mathbf{K}_i)$, dla każdego $1 \leq i \leq n$. Drugie równanie uzyskujemy przez wysumowanie. \square

Siły $\mathbf{F}_i(\mathbf{r}) = \mathbf{F}_i(\mathbf{r}_1, \dots, \mathbf{r}_n)$ nazywamy potencjalnymi, gdy istnieje taka gładka funkcja $V(\mathbf{r})$, że zachodzi

$$\mathbf{F}_i(\mathbf{r}) = -\nabla_{\mathbf{r}_i} V(\mathbf{r}), \quad \text{dla } 1 \leq i \leq n.$$

W takim wypadku równania Newtona (4.2) można zapisać w następującej postaci

$$m_i \ddot{\mathbf{r}}_i = -\nabla_{\mathbf{r}_i} V(\mathbf{r}) + \mathbf{K}_i, \quad 1 \leq i \leq n, \quad (4.3)$$

lub też jako układ równań rzędu pierwszego

$$\begin{cases} \dot{\mathbf{r}}_i = \frac{1}{m_i} \mathbf{p}_i, \\ \dot{\mathbf{p}}_i = -\nabla_{\mathbf{r}_i} V(\mathbf{r}) + \mathbf{K}_i, \end{cases} \quad 1 \leq i \leq n \quad (4.4)$$

Funkcję

$$E(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^n \frac{\|\mathbf{p}_i\|^2}{2m_i} + V(\mathbf{r}), \quad (4.5)$$

gdzie $\mathbf{p} := (\mathbf{p}_1, \dots, \mathbf{p}_n)$, nazywamy energią wewnętrzną układu punktów. Pokażemy następujący fakt.

Lemat 4 (Zmiana energii w czasie). *Jeżeli (\mathbf{r}, \mathbf{p}) spełnia równania (4.4), to*

$$\frac{d}{dt} E(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^n \mathbf{K}_i \cdot \dot{\mathbf{r}}_i.$$

Dowód. Mamy:

$$\begin{aligned} \frac{d}{dt} E(\mathbf{r}, \mathbf{p}) &= \sum_{i=1}^n \left[\nabla_{\mathbf{r}_i} E(\mathbf{r}, \mathbf{p}) \cdot \dot{\mathbf{r}}_i + \nabla_{\mathbf{p}_i} E(\mathbf{r}, \mathbf{p}) \cdot \dot{\mathbf{p}}_i \right] \\ &= \sum_{i=1}^n \left[\nabla_{\mathbf{r}_i} V(\mathbf{r}) \cdot \frac{\mathbf{p}_i}{m_i} + (-\nabla_{\mathbf{r}_i} V(\mathbf{r}) + \mathbf{K}_i) \cdot \frac{\mathbf{p}_i}{m_i} \right] \\ &= \sum_{i=1}^n \mathbf{K}_i \cdot \frac{\mathbf{p}_i}{m_i} = \sum_{i=1}^n \mathbf{K}_i \cdot \dot{\mathbf{r}}_i. \end{aligned}$$

\square

Wniosek 4 (Prawo zachowania energii). *W układzie izolowanym energia jest zachowywana.*

Założmy, że siły wewnętrzne $\mathbf{F}_{k,l}$ oddziaływania pomiędzy punktami k i l , mają kierunek wektora $\mathbf{r}_{k,l} := \mathbf{r}_k - \mathbf{r}_l$, oraz że zależą tylko od ich względnej odległości $r_{k,l} = \|\mathbf{r}_{k,l}\|$. Oznacza to, że

$$\mathbf{F}_{k,l} = F_{k,l}(r_{k,l}) \frac{\mathbf{r}_{k,l}}{r_{k,l}}, \quad \text{dla } 1 \leq k \neq l \leq n, \quad (4.6)$$

gdzie $F_{k,l}$ są gładkimi funkcjami skalarnymi. Siły te są potencjalne. Aby to pokazać, położmy:

$$V_{k,l}(r_{k,l}) := - \int^{r_{k,l}} F_{k,l}(x) dx. \quad (4.7)$$

W powyższym wzorze dolna granica całkowania jest ustalona dowolnie. Mamy:

$$\nabla_{\mathbf{r}_k} V_{k,l}(r_{k,l}) = V'_{k,l}(r_{k,l}) \nabla_{\mathbf{r}_k} r_{k,l} = -F_{k,l}(r_{k,l}) \frac{\mathbf{r}_{k,l}}{r_{k,l}} = -\mathbf{F}_{k,l},$$

gdź

$$\nabla_{\mathbf{r}_k} r_{k,l} = \frac{\mathbf{r}_{k,l}}{r_{k,l}}.$$

Analogicznie pokazujemy, że

$$\nabla_{\mathbf{r}_l} V_{k,l}(r_{k,l}) = \mathbf{F}_{k,l} = -\mathbf{F}_{l,k}.$$

Z powyższych wzorów wynika, że

$$\mathbf{F}_k(\mathbf{r}) := \sum_{\substack{l=1 \\ l \neq k}}^n \mathbf{F}_{k,l} = -\nabla_{\mathbf{r}_l} V(\mathbf{r}), \quad (4.8)$$

gdzie

$$V(\mathbf{r}) = \sum_{1 \leq k \leq l \leq n} V_{k,l}(\mathbf{r}) \quad (4.9)$$

Oznaczmy przez M_k^w moment sił wewnętrznych działających na punkt k , tzn.

$$M_k^w := \mathbf{r}_k \times \mathbf{F}_k(\mathbf{r}) = \mathbf{r}_k \times \sum_{\substack{l=1 \\ l \neq k}}^n \mathbf{F}_{k,l}$$

Analogicznie

$$M_k^z := \mathbf{r}_k \times \mathbf{K}_k,$$

oznacza moment sił zewnętrznych działających na k -ty punkt układu.

Lemat 5. Jeżeli w układzie punkty oddziałują pomiędzy sobą siłami postaci (4.6), to całkowity moment tych sił znika.

Dowód. Mamy pokazać, że

$$M^w = \sum_{k=1}^n M_k^w = \sum_{k=1}^n \mathbf{r}_k \times \sum_{\substack{l=1 \\ l \neq k}}^n \mathbf{F}_{k,l} = \mathbf{0}.$$

Zauważmy, że

$$M^w = \sum_{k \neq l} M_{k,l}^w = \sum_{k < l} (M_{k,l}^w + M_{l,k}^w) \quad \text{gdzie} \quad M_{k,l}^w = \mathbf{r}_k \times \mathbf{F}_{k,l}.$$

Ponadto mamy

$$M_{k,l}^w + M_{l,k}^w = \mathbf{r}_k \times \mathbf{F}_{k,l} + \mathbf{r}_l \times \mathbf{F}_{l,k} = (\mathbf{r}_k - \mathbf{r}_l) \times \mathbf{F}_{k,l} = \mathbf{0},$$

gdź siła $\mathbf{F}_{k,l}$ jest równoległa do wektora $(\mathbf{r}_k - \mathbf{r}_l)$, zobacz (4.6). To kończy dowód. \square

Wniosek 5 (Zasada zachowania momentu pędu). *Jeżeli całkowity moment zewnętrznych sił działających na układ znika*

$$\mathbf{M}^z := \sum_{k=1}^n \mathbf{M}_k^z = \mathbf{0},$$

oraz punkty układu oddziałują pomiędzy sobą siłami postaci (4.6), to całkowity moment pędu układu \mathbf{c} jest stały.

Dowód. Całkowity moment sił wynosi $\mathbf{M} = \mathbf{M}^w + \mathbf{M}^z = \mathbf{0}$. Zatem, z lematu 3 otrzymujemy $\dot{\mathbf{c}} = \mathbf{0}$. \square

Ważnym przykładem sił postaci (4.6) są siły grawitacyjne dla których

$$\mathbf{F}_{k,l} = -\frac{Gm_k m_l}{r_{k,l}^2} \frac{\mathbf{r}_k - \mathbf{r}_l}{r_{k,l}}, \quad V_{k,l} = -\frac{Gm_k m_l}{r_{k,l}}, \quad k \neq l. \quad (4.10)$$

5. Przekształcenia Galileusza i zasada względności

Zasada względności Galileusza odgrywa fundamentalną rolę w fizyce klasycznej. Często wypowiada się ją w bardzo obrazowej i kategorycznej formie mówiąc: we wszystkich układach inercjalnych prawa fizyki są takie same. Warto temu stwierdzeniu nadać bardziej precyzyjny charakter. Jeżeli przez „prawa fizyki” rozumieć równania ruchu, to zasada względności Galileusza mówi, że równania te są takie same we wszystkich układach inercjalnych. Mówiąc precyzyjniej: równania ruchu są niezmiennicze względem transformacji przekształcających jeden układ inercjalny w drugi. Takie transformacje to przekształcenia Galileusza.

§ 8. Przekształcenia Galileusza. Przypomnijmy co to są przekształcenia afiniczne. Niech \mathcal{A} będzie n wymiarową przestrzenią afiniczną, a V_n stowarzyszoną z nią przestrzenią wektorową.

Definicja 2. Przekształcenie $F : \mathcal{A} \rightarrow \mathcal{A}$ nazywamy afinicznym gdy istnieje takie odwzorowanie liniowe $L : V_n \rightarrow V_n$ dla dowolnych $P, Q \in \mathcal{A}$ zachodzi:

$$\tilde{F}(P)F(Q) = L\left(\tilde{F}Q\right). \quad (5.1)$$

Jeżeli OA jest ustalonym punktem, to dowolny $P \in \mathcal{A}$ można zapisać jako $P = O + \mathbf{x}$, gdzie $\mathbf{x} \in V_n$. Położmy $F(P) = O + \mathbf{y}$. Można pokazać, że przekształcenie F jest afiniczne gdy $\mathbf{y} = L(\mathbf{x}) + \mathbf{b}$, gdzie $\mathbf{b} \in V_n$ jest zdefiniowane przez $F(O) = O + \mathbf{b}$.

Zajmiemy się teraz afinicznymi przekształceniami czasoprzestrzeni Newtona $\mathbb{E}^3 \times \mathbb{R}$. Każde takie przekształcenie jest jednoznacznie określone przez odwzorowanie

$$g : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}, \quad (5.2)$$

$$(\mathbf{r}, t) \mapsto (\mathbf{r}', t') = L((\mathbf{r}, t)) + (\mathbf{b}, c), \quad (5.3)$$

gdzie L jest odwzorowaniem liniowym, $\mathbf{b} \in \mathbb{R}^3$ oraz $c \in \mathbb{R}$.

Definicja 3. Odwzorowanie afiniczne $g : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}$ nazywamy przekształceniem Galileusza, dla dowolnych dwóch zdarzeń $(\mathbf{r}_i, t_i), (\mathbf{r}'_i, t'_i) = g(\mathbf{r}_i, t_i)$, $i = 1, 2$, spełnione są warunki:

1. $|t_1 - t_2| = |t'_1 - t'_2|$,
2. jeżeli $t_1 = t_2$, to $\|\mathbf{r}_1 - \mathbf{r}_2\| = \|\mathbf{r}'_1 - \mathbf{r}'_2\|$.

Pierwszy z warunków mówi, że transformacje Galileusza zachowują odstępy czasowe pomiędzy zdarzeniami. Drugi zaś oznacza, że dla zdarzeń równoczesnych jest ono izometrią przestrzeni \mathbb{R}^3 .

Symbolem $O(3, \mathbb{R})$ oznaczamy grupę ortogonalnych, rzeczywistych macierzy wymiaru 3×3 . Analogicznie $SO(3, \mathbb{R})$ stanowi jej podgrupę złożoną z macierzy o wyznaczniku równym 1.

Lemat 6. Dowolne przekształcenie Galileusza $g : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}$ można przedstawić jako złożenie $g = g_1 \circ g_2 \circ g_3$ trzech przekształceń Galileusza $g_i : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}$, dla $i = 1, 2, 3$, następującej postaci:

1. $g_1(\mathbf{r}, t) = (\mathbf{r} + \mathbf{v}t, t)$ gdzie $\mathbf{v} \in \mathbb{R}^3$,
2. $g_2(\mathbf{r}, t) = (\mathbf{r} + \mathbf{b}, t + c)$, gdzie $\mathbf{b} \in \mathbb{R}^3, c \in \mathbb{R}$,
3. $g_3(\mathbf{r}, t) = (A\mathbf{r}, t)$, gdzie $A \in O(3, \mathbb{R})$.

Dowód. Przekształcenia g_1, g_2 oraz g_3 są przekształceniami Galileusza, co wynika z ich definicji.

Weźmy teraz teraz dowolne przekształcenie afiniczne $g : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}$. Zapiszmy je w postaci macierzowej

$$\begin{bmatrix} \mathbf{r}' \\ t' \end{bmatrix} = \begin{bmatrix} A & \mathbf{v} \\ \mathbf{u}^T & \lambda \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ t \end{bmatrix} + \begin{bmatrix} \mathbf{b} \\ c \end{bmatrix}, \quad (5.4)$$

gdzie $A \in \mathbb{M}(3, \mathbb{R})$, $\mathbf{b}, \mathbf{u}, \mathbf{v} \in \mathbb{R}$. W powyższym wzorze wszystkie trójwymiarowe wektory traktujemy jak macierze 3×1 . Zatem

$$\mathbf{r}' = A\mathbf{r} + \mathbf{v}t + \mathbf{b}, \quad t' = \lambda t + \mathbf{u}^T \mathbf{v} + c. \quad (5.5)$$

Założmy, że jest to przekształcenie Galileusza i rozpatrzmy dwa dowolne zdarzenia

$$\mathbf{r}'_i = A\mathbf{r}_i + \mathbf{v}t_i + \mathbf{b}, \quad t'_i = \lambda t_i + \mathbf{u}^T \mathbf{v} + c, \quad i = 1, 2. \quad (5.6)$$

Założmy najpierw, że są one jednoczesne, tzn., że $t_1 = t_2$. Wówczas wówczas $\mathbf{r}' = A\mathbf{r}$, gdzie $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. Z drugiego warunku definicji 3 otrzymujemy $\|\mathbf{r}'\| = \|A\mathbf{r}\| = \|\mathbf{r}\|$, co oznacza, że $A \in O(3, \mathbb{R})$. Poza tym mamy również $0 = t_2 - t_1 = \mathbf{u}^T \mathbf{r}$, dla dowolnego \mathbf{r} . Zatem $\mathbf{u} = \mathbf{0}$.

Pierwszy warunek definicji 3 daje teraz

$$|t'_2 - t'_1| = |\lambda(t_2 - t_1)| = |\lambda|(t_2 - t_1) = |(t_2 - t_1)|,$$

co oznacza, że $\lambda = \pm 1$.

Ostatecznie otrzymujemy, że jeżeli g jest transformacją Galileusza, to jest ona postaci

$$\mathbf{r}' = A\mathbf{r} + \mathbf{v}t + \mathbf{b}, \quad t' = \pm t + \alpha.$$

co dowodzi tezy lematu. □

Powyższy lemat pokazuje, że dowolne przekształcenie Galileusza zależy od parametrów $(A, \mathbf{v}, \mathbf{b}, c, \lambda)$ i ma postać

$$g(\mathbf{r}, t) = \begin{bmatrix} A & \mathbf{v} \\ \mathbf{0}^T & \lambda \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ t \end{bmatrix} + \begin{bmatrix} \mathbf{b} \\ c \end{bmatrix}. \quad (5.7)$$

Takie przekształcenie Galileusza będziemy oznaczać $g(A, \mathbf{v}, \mathbf{b}, c, \lambda)$ gdy będzie konieczne podać jawną zależność od parametrów. Zbiór wszystkich przekształceń Galileusza oznaczymy przez G .

Dla ustalonego $\lambda \in \{-1, +1\}$, zbiór wszystkich przekształceń Galileusza zależy od 10 parametrów. Mamy siedem parametrów $(\mathbf{v}, \mathbf{b}, c) \in \mathbb{R}^7$, oraz 3 parametry pochodzące od $A \in O(3, \mathbb{R})$, gdyż ta grupa Lie jest trójwymiarowa.

Niech $g_i = g(A_i, \mathbf{v}_i, \mathbf{b}_i, c_i, \lambda_i)$, dla $i = 1, 2, 3$. Załóżmy, że zachodzi relacja $g_3 = g_2 \circ g_1$. Wówczas spełnione są następujące zależności:

$$\begin{aligned} A_3 &= A_2 A_1, & \lambda_3 &= \lambda_2 \lambda_1, & \mathbf{v}_3 &= A_2 \mathbf{v}_1 + \mathbf{v}_2 \lambda_1, \\ \mathbf{b}_3 &= A_2 \mathbf{b}_1 + \mathbf{v}_2 c_1 + \mathbf{b}_2, & c_3 &= \lambda_2 c_1 + c_2 \end{aligned} \quad (5.8)$$

Oznaczmy jeszcze $g_e = g(\text{Id}_3, \mathbf{0}, \mathbf{0}, 0, 1)$.

Lemat 7. *Zbiór wszystkich przekształceń Galileusza \mathcal{G} tworzy grupę. Działaniem w tej grupie jest składanie przekształceń, a elementem neutralnym jest g_e .*

Dowód. Formuły (§ 8) pokazują, że zbiór \mathcal{G} jest zamknięty ze względu na składanie odwzorowań. Wynika z nich również, że zachodzą równości $g_e \circ g = g \circ g_e = g$.

Niech $g = g(A, \mathbf{v}, \mathbf{b}, c, \lambda) \in \mathcal{G}$. Sprawdźmy czy istnieje taki element $g' = g(A', \mathbf{v}', \mathbf{b}', c', \lambda') \in \mathcal{G}$, że $g' \circ g = g_e$. Ze wzorów (§ 8) uzyskujemy następujący układ równań

$$\begin{aligned} \text{Id}_3 &= A' A, & 1 &= \lambda' \lambda, & \mathbf{0} &= A' \mathbf{v} + \mathbf{v}' \lambda, \\ \mathbf{0} &= A' \mathbf{b} + \mathbf{v}' c + \mathbf{b}', & 0 &= \lambda' c + c', \end{aligned}$$

który ma dokładnie jedno rozwiązanie

$$\begin{aligned} A' &= A^{-1}, & \lambda' &= \frac{1}{\lambda}, & \mathbf{v}' &= -\frac{1}{\lambda} A^{-1} \mathbf{v}, \\ \mathbf{b}' &= A^{-1} \left(\frac{1}{\lambda} \mathbf{v} - \mathbf{b} \right), & c' &= -\frac{c}{\lambda}. \end{aligned}$$

Pokazuje to, że $g' \in \mathcal{G}$. □

Ważną podgrupą grupy \mathcal{G} stanowią te jej elementy $g = g(A, \mathbf{v}, \mathbf{b}, c, \lambda)$ dla których $\lambda = 1$ i $A \in \text{SO}(3, \mathbb{R})$. Jest ona oznaczana jako \mathcal{G}_1^+ .

§ 9. Zasada względności Galileusza. Zasada ta mówi, że równania ruchu Newtona w układzie inercyjnym są niezmiennicze względem transformacji Galileusza. Nakłada ona bardzo silne ograniczenia na postać sił.

Rozpatrzmy równania ruchu Newtona układu n punktów materialnych:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i(\mathbf{r}, \dot{\mathbf{r}}, t), \quad 1 \leq i \leq n. \quad (5.9)$$

Dowolne przesunięcie w czasie $t \mapsto t + c$ jest transformacją Galileusza. Niezmienniczość równań (5.9) względem nich oznacza, że siły \mathbf{F}_i nie zależą od czasu.

Niezmienniczość równań (5.9) względem transformacji Galileusza $\mathbf{r} \mapsto A\mathbf{r}$, oznacza, że

$$\mathbf{F}_i(A\mathbf{r}, A\dot{\mathbf{r}}) = A\mathbf{F}_i(\mathbf{r}, \dot{\mathbf{r}}) \quad \text{dla } 1 \leq i \leq n.$$

Niezmienniczość względem przesunięć przestrzennych:

$$\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_n) \mapsto (\mathbf{r}_1 + \mathbf{b}, \dots, \mathbf{r}_n + \mathbf{b}),$$

pociąga za sobą to, że siły zależą od względnych położeń $\mathbf{r}_k - \mathbf{r}_l$. Analogicznie, z niezmienniczości względem transformacji Galileusza

$$\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_n) \mapsto (\mathbf{r}_1 + \mathbf{v}t, \dots, \mathbf{r}_n + \mathbf{v}t),$$

wynika, że siły zależą tylko od względnych prędkości $\dot{\mathbf{r}}_k - \dot{\mathbf{r}}_l$.

W powyższych rozważaniach wynika szereg ciekawych wniosków. Jeżeli układ składa się tylko z jednego punktu

$$m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t),$$

to niezmienniczość ze względu na translacje w czasie, położeniach i prędkościach oznacza, że \mathbf{F} . Niezmienniczość względem obrotów oznacza, że $A\mathbf{F} = \mathbf{F}$ dla dowolnego $A \in O(3, \mathbb{R})$, oznacza, że $\mathbf{F} = \mathbf{0}$.

Rozważmy układ składający się z dwóch punktów

$$m_1 \ddot{\mathbf{r}}_1 = \mathbf{F}_1(\mathbf{r}_{1,2}, \dot{\mathbf{r}}_{1,2}), \quad m_2 \ddot{\mathbf{r}}_2 = \mathbf{F}_2(\mathbf{r}_{1,2}, \dot{\mathbf{r}}_{1,2}), \quad (5.10)$$

Zakładamy, że spełnione jest trzecie prawo Newtona. Można pokazać, że przy dowolnych warunkach początkowych zawsze istnieje taki układ inercjalny, w którym oba punkty poruszają się w jednej nieruchomej płaszczyźnie.

Łatwo można sprawdzić, że równania Newtona grawitacyjnego zagadnienia n ciał:

$$m_i \ddot{\mathbf{r}}_i = - \sum_{\substack{j=1 \\ j \neq i}}^n \frac{Gm_i m_j}{\|\mathbf{r}_i - \mathbf{r}_j\|^3} (\mathbf{r}_i - \mathbf{r}_j), \quad 1 \leq i \leq n,$$

są niezmiennicze względem pełnej grupy Galileusza \mathcal{G} .

6. Zagadnienie wielu ciał

Ustalmy pewien układ inercjalny oraz rozważmy izolowany układ n punktów materialnych oddziałujących ze sobą grawitacyjnie. Definiujemy przestrzeń konfiguracyjną wykluczając z \mathbb{R}^{3n} zderzenia \mathcal{C} , gdzie

$$\mathcal{C} = \{\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{R}^{3n} : \mathbf{r}_i = \mathbf{r}_j, i \neq j\}.$$

Układ równań Newtona ma następującą postać:

$$m_i \ddot{\mathbf{r}}_i = - \sum_{\substack{j=1 \\ j \neq i}}^n \frac{Gm_i m_j}{r_{ij}^3} (\mathbf{r}_i - \mathbf{r}_j), \quad 1 \leq i \leq n, \quad (6.1)$$

gdzie, tak jak wcześniej, $r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$.

Równania te można również zapisać jako

$$m_i \ddot{\mathbf{r}}_i = - \frac{\partial V}{\partial \mathbf{r}_i}(\mathbf{r}), \quad 1 \leq i \leq n,$$

gdzie potencjał $V(\mathbf{r}) = V(\mathbf{r}_1, \dots, \mathbf{r}_n)$ określony jest następującym wzorem

$$V(\mathbf{r}) = - \sum_{1 \leq i < j \leq n} \frac{Gm_i m_j}{r_{ij}}.$$

Wprowadźmy następującą diagonalną macierz

$$M := \text{diag}(m_1, m_1, m_1, \dots, m_n, m_n, m_n).$$

Za jej pomocą równanie Newtona można zapisać w zwartej postaci

$$M \ddot{\mathbf{r}} = - \nabla V(\mathbf{r}) \quad (6.2)$$

Wprowadźmy również funkcję Hamiltona dla zagadnienia n ciał. Ma ona postać

$$H(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^n \frac{\|\mathbf{p}_i\|^2}{2m_i} - V(\mathbf{r}). \quad (6.3)$$

Wykorzystując ją możemy równania Newtona zapisać jako układ równań pierwszego rzędu

$$\begin{cases} \dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i} \\ \dot{\mathbf{p}}_i = - \frac{\partial H}{\partial \mathbf{r}_i} = - \frac{\partial V}{\partial \mathbf{r}_i}, \end{cases} \quad (6.4)$$

UWAGA 6.1 Często zamiast potencjału $V(\mathbf{r})$ używa się funkcji $U(\mathbf{r}) = -V(\mathbf{r})$, która nazywa się funkcją siłową.

Przez $\Omega = (\mathbb{R}^{3n} \setminus \mathcal{C}) \times \mathbb{R}^{3n}$ oznaczamy przestrzeń fazową układu.

Lemat 8. Układ (6.4) ma następujące całki pierwsze:

— hamiltonian $H \in C^\infty(\Omega, \mathbb{R})$,

— składowe całkowitego pędu układu $P_1, P_2, P_3 \in C^\infty(\Omega, \mathbb{R})$ zdefiniowane wzorem

$$\mathbf{P}(\mathbf{r}, \mathbf{p}) = (P_1(\mathbf{r}, \mathbf{p}), P_2(\mathbf{r}, \mathbf{p}), P_3(\mathbf{r}, \mathbf{p})) = \mathbf{p}_1 + \cdots + \mathbf{p}_n,$$

— składowe całkowitego momentu pędu układu $C_1, C_2, C_3 \in C^\infty(\Omega, \mathbb{R})$ zdefiniowane wzorem

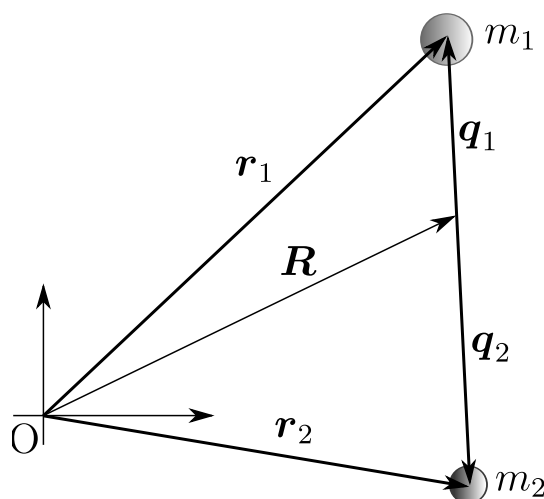
$$\mathbf{C}(\mathbf{r}, \mathbf{p}) = (C_1(\mathbf{r}, \mathbf{p}), C_2(\mathbf{r}, \mathbf{p}), C_3(\mathbf{r}, \mathbf{p})) = \mathbf{r}_1 \times \mathbf{p}_1 + \cdots + \mathbf{r}_n \times \mathbf{p}_n.$$

Poza tym układ posiada całki pierwsze zależne od czasu. Są to współrzędne środka masy; $R_1, R_2, R_3 \in C^\infty(\Omega \times \mathbb{R}, \mathbb{R})$ zdefiniowane wzorem

$$\mathbf{R}(\mathbf{r}, \mathbf{p}, t) = m_1 \mathbf{q}_1 + \cdots + m_n \mathbf{q}_n - \frac{1}{m} \mathbf{P}t.$$

Powyższy lemat nie wyklucza istnienia innych całek pierwszy. W następnym punkcie pokażemy, że dla $n = 2$ takie dodatkowe całki istnieją. Ale przez około dwa wieki nie było wiadomo czy dla $n = 3$ istnieje przynajmniej jedna dodatkowa całka pierwsza. Dopiero całkiem niedawno udało się dowieść, że dla dowolnych dodatnich mass taka całka nie istnieje.

§ 10. Grawitacyjne zagadnienie 2 ciał. Ustalmy pewien układ inercjalny oraz rozważmy izolowany układ dwóch punktów materialnych oddziałujących ze sobą grawitacyjnie. Definiujemy przestrzeń konfiguracyjną wykluczając z $\mathbb{R}^3 \times \mathbb{R}^3$ zderzenia \mathcal{C} , gdzie $\mathcal{C} = \{(\mathbf{r}_1, \mathbf{r}_2) \in \mathbb{R}^3 \times \mathbb{R}^3 : \mathbf{r}_1 = \mathbf{r}_2\}$. Układ równań Newtona ma następującą postać:



Rysunek 1. Grawitacyjne zagadnienie dwóch ciał

$$\begin{cases} m_1 \ddot{\mathbf{r}}_1 = -G \frac{m_1 m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3} (\mathbf{r}_1 - \mathbf{r}_2), \\ m_2 \ddot{\mathbf{r}}_2 = -G \frac{m_1 m_2}{\|\mathbf{r}_2 - \mathbf{r}_1\|^3} (\mathbf{r}_2 - \mathbf{r}_1). \end{cases} \quad (6.5)$$

Zauważmy, że siła grawitacji zależy tylko od względnego wektora wodzącego $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. Korzystając z równań Newtona (6.5) znajdujemy

$$\ddot{\mathbf{r}} = \ddot{\mathbf{r}}_1 - \ddot{\mathbf{r}}_2 = -G \frac{m_1 m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3} (\mathbf{r}_1 - \mathbf{r}_2) + G \frac{m_1 m_2}{\|\mathbf{r}_2 - \mathbf{r}_1\|^3} (\mathbf{r}_2 - \mathbf{r}_1) = \quad (6.6)$$

$$= -G \frac{(m_1 + m_2)}{\|\mathbf{r}\|^3} \mathbf{r} = -\frac{\mu}{r^3} \mathbf{r}, \quad (6.7)$$

gdzie $\mu = G(m_1 + m_2)$ oraz $\|\mathbf{r}\| = r$. Zatem zamiana zmiennych $(\mathbf{r}_1, \mathbf{r}_2) \mapsto (\mathbf{r}, \mathbf{r}_2)$ przekształca równania (6.5) do postaci

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3} \mathbf{r}, \quad (6.8)$$

$$\ddot{\mathbf{r}}_2 = \frac{Gm_1}{r^3} \mathbf{r}. \quad (6.9)$$

W ten sposób pierwsze z powyższych równań można rozwiązywać niezależnie od drugiego.

Innym sposobem uzyskania separacji równań ruchu jest przejście do współrzędnych środka masy. Niech $\mathbf{r}_i = \mathbf{R} + \mathbf{q}_i$, dla $i = 1, 2$, gdzie

$$\mathbf{R} = \frac{1}{m_1 + m_2} (m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2).$$

Wektory $\mathbf{q}_1, \mathbf{q}_2$ spełniają dokładnie takie same równania jak wektory \mathbf{r}_1 i \mathbf{r}_2 (dlaczego?), to znaczy

$$\ddot{\mathbf{q}}_1 = -G \frac{m_2}{\|\mathbf{q}_1 - \mathbf{q}_2\|^3} (\mathbf{q}_1 - \mathbf{q}_2), \quad (6.10)$$

$$\ddot{\mathbf{q}}_2 = -G \frac{m_1}{\|\mathbf{q}_2 - \mathbf{q}_1\|^3} (\mathbf{q}_2 - \mathbf{q}_1). \quad (6.11)$$

Załóżmy, że rozpatrujemy nasze zagadnienie w układzie którego początek znajduje się w środku masy. Wówczas $m_1 \mathbf{q}_1 + m_2 \mathbf{q}_2 = \mathbf{0}$, a zatem $\mathbf{q}_2 = -(m_1/m_2) \mathbf{q}_1$. Podstawiając tę relację do równania (6.10) otrzymujemy

$$\ddot{\mathbf{q}}_1 = -\frac{Gm_2}{\|\mathbf{q}_1 + \frac{m_1}{m_2} \mathbf{q}_1\|^3} \left(1 + \frac{m_1}{m_2}\right) \mathbf{q}_1 = -\frac{\mu_1}{q_1^3} \mathbf{q}_1, \quad (6.12)$$

gdzie

$$\mu_1 = \frac{Gm_2^3}{(m_1 + m_2)^2}, \quad \|\mathbf{q}_1\| = q_1.$$

Podobnie eliminując \mathbf{q}_1 z równania (6.11) uzyskamy

$$\ddot{\mathbf{q}}_2 = -\frac{\mu_2}{q_2^3} \mathbf{q}_2, \quad (6.13)$$

gdzie

$$\mu_2 = \frac{Gm_1^3}{(m_1 + m_2)^2}, \quad \|\mathbf{q}_2\| = q_2.$$

Widzimy zatem, że rozwiązanie problem dwóch ciał sprowadza się do rozwiązywania równania postaci $\ddot{\mathbf{q}} = -(\mu/q^3)\mathbf{q}$. Zapiśmy je jako układ równań pierwszego rzędu postaci

$$\begin{cases} \dot{\mathbf{q}} = \mathbf{v}, \\ \dot{\mathbf{v}} = -\frac{\mu}{q^3}\mathbf{q}. \end{cases} \quad (6.14)$$

Równania (6.14) mają następujące całki pierwsze:

— całkę energii

$$h = \frac{1}{2}\|\mathbf{v}\|^2 - \frac{\mu}{q},$$

— całkę momentu pędu:

$$\mathbf{c} = \mathbf{q} \times \mathbf{v}.$$

UWAGA 6.2 Rozważmy ruch punktu materialnego o jednostkowej masie w polu nieruchomego centrum grawitacyjnego umieszczonego w początku układu współrzędnych. Jest to zagadnienie Keplera. Równania ruchu punktu mają postać (6.14), przy czym \mathbf{q} jest jego promieniem wodzącym, a $\mu = GM$, gdzie M jest masą centrum grawitacyjnego.

Lemat 9. Orbita punktu \mathbf{q} leży w płaszczyźnie przechodzącej przez początek układu współrzędnych i prostopadłej do wektora \mathbf{c} lub odbywa się po prostej przechodzącej przez początek układu współrzędnych.

Dowód. Jeżeli $\mathbf{c} \neq \mathbf{0}$, to $\mathbf{q} \perp \mathbf{c}$, co implikuje tezę.

Aby rozważyć przypadek $\mathbf{c} = \mathbf{0}$, zauważmy, że zachodzi

$$\frac{d}{dt} \left(\frac{\mathbf{q}}{q} \right) = \frac{q\dot{\mathbf{q}} - \dot{q}\mathbf{q}}{q^2} = \frac{q^2\ddot{\mathbf{q}} - \dot{q}q\dot{\mathbf{q}}}{q^3} = \frac{\mathbf{q} \times (\dot{\mathbf{q}} \times \mathbf{q})}{q^3} = \frac{\mathbf{c} \times \mathbf{q}}{q^3}$$

W powyższych rachunkach wykorzystaliśmy następującą tożsamość:

$$\mathbf{a}_1 \times (\mathbf{a}_2 \times \mathbf{a}_3) = (\mathbf{a}_1 \cdot \mathbf{a}_3)\mathbf{a}_2 - (\mathbf{a}_1 \cdot \mathbf{a}_2)\mathbf{a}_3.$$

Zatem, gdy $\mathbf{c} = \mathbf{0}$, to wektor jednostkowy $\frac{\mathbf{q}}{q}$ jest stały, a to oznacza, że ruch punktu odbywa się po prostej. \square

Definicja 4. Wektorem Laplace'a-Rungego-Lentza nazywamy wektor

$$\mathbf{e} = \frac{1}{\mu}\mathbf{v} \times \mathbf{c} - \frac{\mathbf{q}}{q}.$$

Zauważmy, że jeżeli $\mathbf{c} \neq \mathbf{0}$, to $\mathbf{e} \cdot \mathbf{c} = 0$. Zatem, wektor \mathbf{e} znajduje się w płaszczyźnie ruchu punktu materialnego. Natomiast dla $\mathbf{c} = \mathbf{0}$ otrzymujemy, że $\mathbf{e} = -\frac{\mathbf{q}}{q}$, czyli wtedy wektor \mathbf{e} wskazuje na kierunek ruchu punktu materialnego po prostej.

Lemat 10. Współrzędne wektora Laplace'a-Rungego-Lentza są całkami pierwszymi układu (6.14).

Dowód. Mamy

$$\dot{\mathbf{e}} = \frac{1}{\mu} \dot{\mathbf{v}} \times \mathbf{c} + \frac{1}{\mu} \mathbf{v} \times \dot{\mathbf{c}} + \frac{\mathbf{q} \times \mathbf{c}}{q^3} = \frac{1-\mu}{\mu} \frac{\mathbf{q} \times \mathbf{c}}{q^3} = \mathbf{0},$$

gdyż $\dot{\mathbf{c}} = \mathbf{0}$, oraz $\dot{\mathbf{v}} = -(\mu/q^3)\mathbf{q}$, na mocy równań ruchu. \square

Reasumując, dla sześciu równań (6.14) znamy siedem całek pierwszych h , c i e . Wnika stąd, że nie są one funkcjonalnie niezależne, gdyż układ sześciu równań różniczkowych może mieć maksymalnie pięć funkcjonalnie niezależnych całek pierwszych. Pokazaliśmy wcześniej związek pomiędzy całkami e oraz c . Zachodzi mianowicie $e \cdot c = 0$. Pokażemy, że zachodzi jeszcze jeden związek pomiędzy całkami pierwszymi.

Lemat 11. *Spełniony jest następujący związek:*

$$\mu^2 (e^2 - 1) = 2hc^2. \quad (6.15)$$

Dowód. Mamy

$$e^2 = \mathbf{e} \cdot \mathbf{e} = \frac{1}{\mu^2} (\mathbf{v} \times \mathbf{c}) \cdot (\mathbf{v} \times \mathbf{c}) + 1 - 2 \frac{1}{\mu q} (\mathbf{v} \times \mathbf{c}) \cdot \mathbf{q}.$$

Ale

$$(\mathbf{v} \times \mathbf{c}) \cdot (\mathbf{v} \times \mathbf{c}) = v^2 c^2,$$

gdyż wektory \mathbf{v} i \mathbf{c} są prostopadłe do siebie. Dodatkowo mamy

$$(\mathbf{v} \times \mathbf{c}) \cdot \mathbf{q} = (\mathbf{q} \times \mathbf{v}) \cdot \mathbf{c} = c^2.$$

Więc

$$e^2 = \frac{v^2 c^2}{\mu^2} + 1 - 2 \frac{c^2}{\mu q},$$

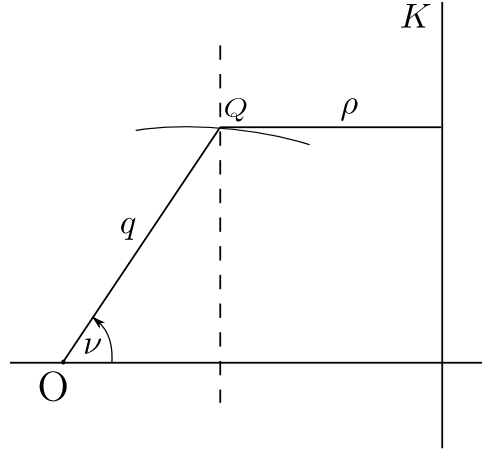
co po prostych przekształceniach daje związek (6.15). \square

Definicja 5. *Krzywą stożkową nazywamy zbiór punktów Q spełniający warunek: stosunek ich odległości q od stałego punktu O , zwanego ogniskiem, do ich odległości ρ od ustalonej prostej K , zwanej kierownicą, jest stały $\frac{q}{\rho} = e$ tzn. stosunek odległości q punktu Q od ogniska O do odległości ρ punktu Q od kierownicy k jest stały. Wielkość e nosi nazwę mimośrodu krzywej stożkowej.*

Niech d oznacza odległość ogniska do kierownicy. Z rysunku 2 łatwo widać, że $d = q \cos \nu + \rho$. Z definicji mamy, że $q = \rho e$. W konsekwencji otrzymujemy

$$q = \frac{de}{1 + e \cos \nu}. \quad (6.16)$$

Zatem wszystkie krzywe stożkowe można opisać równaniem (??) we współrzędnych biegunowych, gdzie q, ν to współrzędne punktu, zaś e to mimośród krzywej decydujący o kształcie (dla $0 \leq e < 1$ mamy elipsę, ze szczególnym przypadkiem okręgu dla $e = 0$, dla $e = 1$ mamy parabolę a dla $e > 1$ hiperbolę).



Rysunek 2. Kierownica o ognisko stożkowej

Przejdźmy teraz do interpretacji geometrycznej wektora e . Mnożąc wektor e przez q otrzymujemy

$$e \cdot q = \frac{1}{\mu}(\mathbf{v} \times \mathbf{c}) \cdot \mathbf{q} - \frac{1}{q} \mathbf{q} \cdot \mathbf{q}, \quad (6.17)$$

$$e \cdot q = \frac{c^2}{\mu} - q, \quad (6.18)$$

$$e \cdot q + q = \frac{c^2}{\mu}. \quad (6.19)$$

W przypadku, gdy $e = 0$, to $q = c^2/\mu$. Zatem w tym przypadku długość wektora q jest stała, inaczej mówiąc orbita ruchu punktu materialnego jest okręgiem. Jeżeli $e \neq 0$, to oznaczmy kąt pomiędzy wektorem e , a wektorem wodzącym q , przez ν . W mechanice nieba nazywa się on anomalią prawdziwą. Ponieważ $e \cdot q = eq \cos \nu$, to z równania (6.19) otrzymujemy

$$q = q(\nu) = \frac{p}{1 + e \cos \nu}, \quad p = \frac{c^2}{\mu}. \quad (6.20)$$

Zatem orbity po jakich porusza się punkt materialny są krzywymi stożkowymi, dlatego też $p = a(1 - e^2)$. Sklasyfikujemy orbity ruchu wykorzystując uzyskane związki:

1. Dla $e = 1$ otrzymujemy, że $a \rightarrow +\infty$. Zatem $h \rightarrow 0$ i mamy ruch po paraboli.
2. Dla $0 \leq e < 1$ otrzymujemy, że $a > 0$. Zatem $h < 0$ i mamy ruch po elipsie.
3. Dla $e > 1$ otrzymujemy, że $a < 0$. Zatem $h > 0$ i mamy ruch po hiperboli.

Podsumowując, w problemie Keplera rozpatrujemy równanie $\ddot{\mathbf{q}} = -(\mu/q^3)\mathbf{q}$. Znane całki ruchu c , e oraz h pozwalają na opisanie kształtu orbity ruchu punktu materialnego w formie toru $q(\nu)$. Ruch odbywa się po krzywych stożkowych. Dalej należy uporać się z problemem jak punkt materialny

porusza się po takiej orbicie. Do tego celu wykorzystamy lemat 9 oraz całkę momentu pędu c zapisaną we współrzędnych biegunowych (q, φ) :

$$\mathbf{c} = \mathbf{q} \times \dot{\mathbf{q}} = q\mathbf{e}_q \times (\dot{q}\mathbf{e}_q + \dot{\varphi}q\mathbf{e}_\varphi) = q\dot{q}\mathbf{e}_q \times \mathbf{e}_q + \dot{\varphi}q^2\mathbf{e}_q \times \mathbf{e}_\varphi = \dot{\varphi}q^2\mathbf{e}_c, \quad (6.21)$$

gdzie $\mathbf{e}_q, \mathbf{e}_\varphi$ są wektorami jednostkowymi we współrzędnych biegunowych, a $\mathbf{e}_c = \frac{\mathbf{c}}{c}$. Zatem pozostaje rozwiązać równanie postaci

$$c = \dot{\varphi}q^2, \quad (6.22)$$

lub równoważnie równanie

$$c = \dot{\nu}q^2. \quad (6.23)$$

W tym celu stosujemy zamianę zmiennych $t \rightarrow t(u)$ postaci

$$\frac{du}{dt} = \frac{\chi}{q},$$

gdzie χ jest stała, którą ustalimy później. Zatem

$$u = \int_{\tau}^t \frac{\chi}{q(s)} ds,$$

gdzie dolną granicę całkowania τ również ustalimy później w zależności od energii. Zauważmy, że

$$\frac{dq}{dt} = \frac{dq}{du} \frac{du}{dt} = \frac{dq}{du} \frac{\chi}{q}.$$

Podstawmy otrzymaną tożsamość do równania energii:

$$\dot{q}^2(t)q^2(t) + c^2 = 2(\mu q(t) + q^2(t)h) \quad (6.24)$$

Uzyskamy nowe równanie postaci

$$\dot{q}^2(u)\chi^2 + c^2 = 2(\mu q(u) + q^2(u)h), \quad (6.25)$$

gdzie \dot{q} stanowi pochodną q po zmiennej u . Dalej rozpatrujemy poszczególne przypadki:

1. Przypadek paraboliczny: $h = 0$, postulujemy $\chi^2 = \mu$ otrzymując równania Bakera, algebraiczne równania trzeciego rzędu ze względu na zmienną u dające się rozwiązać w analityczny sposób:

$$\begin{cases} \frac{1}{6}u^3 + \frac{c^2}{2\mu}u = \sqrt{\mu}(t - \tau) \\ q(u) = \frac{1}{2}u^2 + \frac{c^2}{2\mu} \end{cases} \quad (6.26)$$

2. Przypadek $h \neq 0$ postulujemy, że $\chi^2 = \mu/a$ i otrzymują równania Keplera postaci:

— gdy $h > 0$, przypadek hiperboliczny

$$\begin{cases} \sqrt{\frac{\mu}{a^3}}(t - \tau) = es \sinh u - u \\ q(u) = a(e \cosh u - 1) \end{cases} \quad (6.27)$$

— gdy $h < 0$, przypadek eliptyczny

$$\begin{cases} \sqrt{\frac{\mu}{a^3}}(t - \tau) = u - es \sin u \\ q(u) = a(1 - e \cos u) \end{cases} \quad (6.28)$$

7. Układy z więzami i formalizm Lagrange'a

Rozważmy układ izolowany n punktów materialnych $\mathbf{r}_1, \dots, \mathbf{r}_n$ o masach m_1, \dots, m_n . Zakładamy, że punkty oddziałują pomiędzy sobą. Wypadkowa siła \mathbf{F}_i działająca na i -ty punkt materialny wynosi

$$\mathbf{F}_i = \sum_{\substack{j=1 \\ j \neq i}}^n \mathbf{F}_{i,j}.$$

Dla prostoty zakładamy, że

$$\mathbf{F}_{i,j} = F_{i,j}(r_{i,j}) \frac{\mathbf{r}_i - \mathbf{r}_j}{r_{i,j}},$$

gdzie $r_{i,j} = \|\mathbf{r}_i - \mathbf{r}_j\|$ oraz $F_{i,j}$ jest ciągłą funkcją skalarną zależną od odległości $r_{i,j}$. Zatem istnieje potencjał $V_{i,j}(\mathbf{r})$ taki, że $\mathbf{F}_{ij} = -\nabla_{\mathbf{r}_i} V_{ij}(\mathbf{r})$.

Rozpatrzmy sytuację gdy na ruch punktów nałożone są pewne ograniczenia (więzy). Na przykład, dany punkt może poruszać się po pewnej powierzchni lub krzywej lub też odległość pomiędzy dwoma punktami ma być stała. Ograniczenia tego typu ograniczają liczbę stopni swobody układu, powodują, że pojawia się nowa siła działająca na punkty co w istotny sposób zmienia równania ruchu.

Definicja 6. *Więzy nazywamy holonomicznymi, gdy można je opisać zbiorem równań postaci $f_l(\mathbf{r}_1, \dots, \mathbf{r}_n) = 0$, $l = 1, \dots, K$, dla których*

$$\text{rank} \begin{bmatrix} \frac{\partial f_1}{\partial \mathbf{r}_1} & \dots & \frac{\partial f_1}{\partial \mathbf{r}_n} \\ \vdots & & \vdots \\ \frac{\partial f_K}{\partial \mathbf{r}_1} & \dots & \frac{\partial f_K}{\partial \mathbf{r}_n} \end{bmatrix} = K.$$

Założmy, że więzy definiują rozmaitość gładką M . Niech $\delta \mathbf{r}_1, \dots, \delta \mathbf{r}_n$ będą wektorami stycznymi do M w ustalonym czasie. Założmy, że dla każdego $i = 1, \dots, n$ więzy nałożone na i -ty punkt materialny możemy wyrazić za pomocą nowej siły \mathbf{N}_i działającej na ten punkt, tzw. siły reakcji więzów. Wówczas wypadkowa siła działająca na i -ty punkt wynosi $\mathbf{F}_i + \mathbf{N}_i$.

Równania ruchu możemy zapisać następująco: $m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i + \mathbf{N}_i$. Otrzymujemy zasadę d'Alemberta dla układu punktów:

$$\sum_{i=1}^n (m_i \ddot{\mathbf{r}}_i - \mathbf{F}_i) \cdot \delta \mathbf{r}_i = \sum_{i=1}^n \mathbf{N}_i \cdot \delta \mathbf{r}_i = 0.$$

Z warunku holonomiczności mamy lokalne przedstawienie

$$\mathbf{r}_i = \mathbf{r}_i(\mathbf{q}_1, \dots, \mathbf{q}_{3n-K}),$$

dla $i = 1, \dots, n$. Stąd otrzymujemy, że

$$\begin{aligned} \mathbf{v}_i = \dot{\mathbf{r}}_i &= \sum_{k=1}^{3n-K} \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} \dot{\mathbf{q}}_k \\ \sum_{i=1}^n \mathbf{F}_i \cdot \delta \mathbf{r}_i &= \sum_{k=1}^{3n-K} \underbrace{\left(\sum_{i=1}^n \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} \right)}_{\text{ozn. } Q_k} \delta \mathbf{q}_k = \sum_{k=1}^{3n-K} Q_k \delta \mathbf{q}_k. \end{aligned}$$

Korzystając z potencjalności sił \mathbf{F}_i dostajemy następującą postać Q_k :

$$Q_k = - \sum_{i=1}^n \frac{\partial V}{\partial \mathbf{r}_i} \cdot \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} = - \frac{\partial V}{\partial \mathbf{q}_k}(\mathbf{r}_1(q), \dots, \mathbf{r}_n(q)) = - \frac{\partial V \circ \mathbf{r}}{\partial \mathbf{q}_k}(q),$$

gdzie $V \circ \mathbf{r}$ będziemy oznaczać przez V . Zatem

$$\sum_{k=1}^{3n-K} Q_k \delta \mathbf{q}_k = - \sum_{k=1}^{3n-K} \frac{\partial V}{\partial \mathbf{q}_k} \delta \mathbf{q}_k.$$

Zajmiemy się teraz sumą $\sum_{i=1}^n m_i \ddot{\mathbf{r}}_i \cdot \delta \mathbf{r}_i$:

$$\begin{aligned} \sum_{i=1}^n m_i \ddot{\mathbf{r}}_i \cdot \delta \mathbf{r}_i &= \sum_{k=1}^{3n-K} \left(\sum_{i=1}^n m_i \ddot{\mathbf{r}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} \right) \delta \mathbf{q}_k = \\ &= \sum_{k=1}^{3n-K} \left(\sum_{i=1}^n m_i \left(\frac{d}{dt} \left(\dot{\mathbf{r}}_i \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} \right) - \dot{\mathbf{r}}_i \frac{d}{dt} \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} \right) \right) \delta \mathbf{q}_k. \end{aligned}$$

Zauważmy, że

$$\frac{d}{dt} \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} = \frac{\partial \dot{\mathbf{r}}_i}{\partial \mathbf{q}_k},$$

a zatem

$$\frac{d}{dt} \left(\dot{\mathbf{r}}_i \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} \right) - \dot{\mathbf{r}}_i \frac{d}{dt} \frac{\partial \mathbf{r}_i}{\partial \mathbf{q}_k} = \frac{d}{dt} \left(\frac{\partial}{\partial \dot{\mathbf{q}}_k} \left(\frac{1}{2} \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i \right) \right) - \frac{\partial}{\partial \dot{\mathbf{q}}_k} \left(\frac{1}{2} \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i \right).$$

Uwzględniając powyższe otrzymujemy

$$\begin{aligned} \sum_{k=1}^{3n-K} \left(\frac{d}{dt} \left(\frac{\partial}{\partial \dot{\mathbf{q}}_k} \left(\frac{1}{2} \sum_{i=1}^n m_i \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i \right) \right) - \frac{\partial}{\partial \dot{\mathbf{q}}_k} \left(\frac{1}{2} \sum_{i=1}^n \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i \right) \right) \cdot \delta \mathbf{q}_k \\ = - \sum_{k=1}^{3n-K} \frac{\partial V}{\partial \mathbf{q}_k} \delta \mathbf{q}_k \end{aligned}$$

Ostatecznie równania Lagrange'a mają postać

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0, \quad k = 1, \dots, 3n - K, \quad (7.1)$$

gdzie

$$L(q, \dot{q}) = T(q, \dot{q}) - V(q),$$

$$T(q, \dot{q}) = \frac{1}{2} \sum_{i=1}^n m_i \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i.$$

Funkcję $L(q, \dot{q})$ nazywamy funkcją Lagrange'a.

PRZYKŁAD 7.1 Rozważmy wahadło matematyczne, mamy do czynienia z więzami postaci: $x^2 + y^2 = l^2$. Współrzędne (x, y) możemy wyrazić jako $x = l \cos \varphi$, $y = l \sin \varphi$. Wówczas

$$T(x, y, \dot{x}, \dot{y}) = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2) = \frac{1}{2} m l^2 \dot{\varphi}^2$$

$$V(x, y) = -gmx = -gml \cos \varphi.$$

Zatem $L(x, y, \dot{x}, \dot{y}) = \frac{1}{2} m l^2 \dot{\varphi}^2 - gml \cos \varphi$ a równania mają następującą postać:

$$m l^2 \ddot{\varphi} + gml \sin \varphi = 0 \quad \equiv \quad \ddot{\varphi} = -\omega_0^2 \sin \varphi,$$

gdzie $\omega_0^2 = \frac{g}{l}$.

Prawo 5. (Zasada Hamiltona) Rozważmy układ mechaniczny z N stopniami swobody opisany we współrzędnych uogólnionych $\mathbf{q} = (\mathbf{q}_1, \dots, \mathbf{q}_N)$ z funkcją Lagrange'a $L(q, \dot{q})$. Wówczas

$$\frac{\delta I}{\delta q} = 0,$$

gdzie

$$I[q] = \int_{t_1}^{t_2} L(q(t), \dot{q}(t)) dt.$$

8. Równania Hamiltona

Przypomnijmy, że w formalizmie Lagrange'a równania ruchu mają następującą postać

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_k} - \frac{\partial L}{\partial q_k} = 0 \quad (8.1)$$

dla $k = 1, \dots, n$. Sformułujemy dalej formalizm Hamiltona opisując ruch za pomocą układu $2n$ równań pierwszego rzędu. Z matematycznego punktu widzenia zmiana formalizmu z Lagrange'a na Hamiltona sprowadza się do zmiany zmiennych (q, \dot{q}, t) na (q, p, t) za pomocą transformacji Legendre'a $\dot{q} = \dot{q}(q, p)$.

Definicja 7. Pęd uogólniony (kanoniczny) $p = (p_1, \dots, p_n)$ definiujemy jako

$$p_k = \frac{\partial L}{\partial \dot{q}_k}(q, \dot{q}, t), \quad k = 1, \dots, n. \quad (8.2)$$

Współrzędne (q, p) nazywamy współrzędnymi uogólnionymi (kanonicznymi).

Definicja 8. Hamiltonianem nazywamy funkcję $H : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ klasy C^∞ postaci

$$H = H(q, p) = \sum_{k=1}^n \dot{q}_k(q, p) p_k - L(q, \dot{q}(q, p)). \quad (8.3)$$

Twierdzenie 2. Równania Lagrange'a (8.1) są równoważne następującemu układowi równań Hamiltona

$$\begin{cases} \dot{q}_k = \frac{\partial H}{\partial p_k} \\ \dot{p}_k = -\frac{\partial H}{\partial q_k} \end{cases}, \quad (8.4)$$

dla $k = 1, \dots, n$.

Dowód. Ustalmy $k \in \{1, \dots, n\}$ oraz obliczmy różniczkę H uwzględniając definicję pędu uogólnionego p :

$$\begin{aligned} dH &= \frac{\partial H}{\partial q_k} dq_k + \frac{\partial H}{\partial p_k} dp_k + \frac{\partial H}{\partial t} dt = -\frac{\partial L}{\partial q_k} dq_k + \dot{q}_k dp_k + p_k d\dot{q}_k - \frac{\partial L}{\partial \dot{q}_k} d\dot{q}_k \\ &\quad - \frac{\partial L}{\partial t} dt = -\frac{\partial L}{\partial q_k} dq_k + \dot{q}_k dp_k - \frac{\partial L}{\partial t} dt. \end{aligned}$$

Ponadto z równań Lagrange'a (8.1) wynika, że $\frac{\partial L}{\partial q_k} = \dot{p}_k$ a zatem

$$dH = -\dot{p}_k dq_k + \dot{q}_k dp_k - \frac{\partial L}{\partial t} dt.$$

Ostatecznie, porównując postacie różniczki dH , otrzymujemy, że

$$\frac{\partial H}{\partial q_k} = -\dot{p}_k \quad \text{oraz} \quad \frac{\partial H}{\partial p_k} = \dot{q}_k,$$

co dowodzi tezy twierdzenia. □

PRZYKŁAD 8.1 Rozpatrzmy układ taki, że

$$L(x, y) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - V(x, y).$$

Wówczas

$$p_x = \frac{\partial L}{\partial \dot{x}} = \dot{x}, \quad p_y = \frac{\partial L}{\partial \dot{y}} = \dot{y}.$$

Stąd hamiltonian H dany jest wzorem

$$H = H(x, y, p_x, p_y) = \dot{x} p_x + \dot{y} p_y - L(x, y, \dot{x}, \dot{y}) = \frac{1}{2}(p_x^2 + p_y^2) + V(x, y)$$

a równania Hamiltona mają następującą postać

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial p_x} = p_x \\ \dot{y} = \frac{\partial H}{\partial p_y} = p_y \\ \dot{p}_x = -\frac{\partial H}{\partial x} = -\frac{\partial V}{\partial x} \\ \dot{p}_y = -\frac{\partial H}{\partial y} = -\frac{\partial V}{\partial y} \end{cases}.$$

Niech $z = (q_1, \dots, q_n, p_1, \dots, p_n)$ wówczas równania Hamiltona (8.4) można zapisać w postaci

$$\dot{z} = J \frac{\partial H}{\partial z}, \quad (8.5)$$

gdzie $J = \begin{bmatrix} \Theta_n & \text{Id}_n \\ -\text{Id}_n & \Theta_n \end{bmatrix}$ jest jednostkową macierzą symplektyczną.

Lemat 12. Jednostkowa macierz symplektyczna J posiada następujące własności:

1. $J^2 = -\text{Id}_{2n}$,
2. $J^{-1} = -J$,
3. $\det J = 1$.

9. Przekształcenia kanoniczne

Rozważmy dyfeomorfizm $\Phi : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ taki, że $z = \Phi(\xi)$. Interesują nas tylko takie zamiany zmiennych Φ , dla których $\Phi^{-1}(z)$ są współzrzednymi kanonicznymi, tj. istnieje funkcja $K : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ klasy C^∞ taka, że w nowych współzrzednych równania ruchu mają postać równań Hamiltona $\dot{\xi} = J \frac{\partial K}{\partial \xi}$. Funkcja K stanowi hamiltonian w zbiorze nowych współzrzednych. Różniczkując równanie definiujące zamianę zmiennych po t otrzymujemy, że $\dot{z} = \frac{\partial \Phi}{\partial \xi} \dot{\xi}$, a zatem

$$\dot{\xi} = \left(\frac{\partial \Phi}{\partial \xi} \right)^{-1} J \frac{\partial H}{\partial z}(\Phi(\xi)),$$

lub równoważnie

$$\dot{\xi} = \left(\frac{\partial \Phi}{\partial \xi} \right)^{-1} J \left(\frac{\partial \Phi}{\partial \xi} \right)^{-T} \frac{\partial H}{\partial z},$$

ponieważ

$$\frac{\partial H}{\partial z} = \left(\frac{\partial \Phi}{\partial \xi} \right)^{-T} \frac{\partial H}{\partial \xi}.$$

Ostatecznie nowe równania ruchu mają postać równań Hamiltona wtedy i tylko wtedy, gdy

$$-J \left(\frac{\partial \Phi}{\partial \xi} \right)^{-1} J \left(\frac{\partial \Phi}{\partial \xi} \right)^{-T} \frac{\partial H}{\partial \xi} = \frac{\partial K}{\partial \xi}.$$

Dla $K(\xi) = H(\Phi(\xi))$ otrzymujemy warunek postaci

$$\left(\frac{\partial\Phi}{\partial\xi}\right)^{-1} J \left(\frac{\partial\Phi}{\partial\xi}\right)^{-T} = J$$

lub równoważnie

$$J = \frac{\partial\Phi}{\partial\xi} J \left(\frac{\partial\Phi}{\partial\xi}\right)^T.$$

Zauważmy, że warunek na Φ ma równoważną postać

$$\left(\frac{\partial\Phi}{\partial\xi}\right)^T J \left(\frac{\partial\Phi}{\partial\xi}\right) = J,$$

co wynika z analogicznego wnioskowania. Rzeczywiście, przekształcając równanie $\dot{z} = J \frac{\partial H}{\partial z}$ otrzymujemy

$$\frac{\partial\Phi}{\partial\xi} \dot{\xi} = J \left(\frac{\partial\Phi}{\partial\xi}\right)^{-T} \frac{\partial K}{\partial \xi},$$

co jest dalej równoważne

$$\left(\frac{\partial\Phi}{\partial\xi}\right)^T J^{-1} \frac{\partial\Phi}{\partial\xi} \dot{\xi} = \frac{\partial K}{\partial \xi},$$

co dowodzi, że

$$\left(\frac{\partial\Phi}{\partial\xi}\right)^T J^{-1} \frac{\partial\Phi}{\partial\xi} = J^{-1}.$$

Definicja 9. Dyfeomorfizm $\Phi : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ nazywamy przekształceniem kanonicznym (symplektycznym), gdy

$$\frac{\partial\Phi}{\partial\xi} J \left(\frac{\partial\Phi}{\partial\xi}\right)^T = J.$$

Lemat 13. Zbiór przekształceń symplektycznych z operacją składania odwzorowań tworzy grupę $\text{Sp}(2n, \mathbb{R})$.

Dowód. Dla dowodu tezy lematu wystarczy pokazać, że zbiór przekształceń symplektycznych jest zamknięty ze względu na składanie i branie elementu odwrotnego. Rzeczywiście ustalmy dowolne dwa przekształcenia symplektyczne Φ_1 oraz Φ_2 , gdzie $\Phi_1(\xi_1) = \xi_2$ oraz $\Phi_2(\xi_2) = z$. Wówczas dla przekształcenia $\Phi_3 = \Phi_2 \circ \Phi_1$ z symplektyczności odwzorowań Φ_1 i Φ_2 otrzymujemy, że

$$\frac{\partial\Phi_3}{\partial\xi_1} J \left(\frac{\partial\Phi_3}{\partial\xi_1}\right)^T = \frac{\partial\Phi_2}{\partial\xi_2} \frac{\partial\Phi_1}{\partial\xi_1} J \left(\frac{\partial\Phi_1}{\partial\xi_1}\right)^T \left(\frac{\partial\Phi_2}{\partial\xi_2}\right)^T = \frac{\partial\Phi_2}{\partial\xi_2} J \left(\frac{\partial\Phi_2}{\partial\xi_2}\right)^T = J.$$

Ponadto

$$\frac{\partial\Phi_1^{-1}}{\partial\xi_2} J \left(\frac{\partial\Phi_1^{-1}}{\partial\xi_2}\right)^T = \left(\frac{\partial\Phi_1}{\partial\xi_1}\right)^{-1} J \left(\frac{\partial\Phi_1}{\partial\xi_1}\right)^{-T} = J,$$

co dowodzi tezy lematu. □

10. Wektorowe przestrzenie symplektyczne

Definicja 10. Symplektyczną przestrzenią wektorową nazywamy dowolną przestrzeń wektorową V nad \mathbb{R} wyposażoną w formę dwuliniową $\omega : V \times V \rightarrow \mathbb{R}$ spełniającą następujące warunki

- (1) $\omega(u, v) = -\omega(v, u)$ dla każdego $u, v \in V$ (antysymetryczność),
- (2) $\omega(u, v) = 0$ dla każdego $v \in V$, to $u = 0$ (niezdegenerowanie),

Formę ω nazywamy formą symplektyczną.

Definicja 11. Niech (V, ω) będzie symplektyczną przestrzenią wektorową. Przekształcenie liniowe $A : V \rightarrow V$ nazywamy przekształceniem symplektycznym, o ile $\omega(Au, Av) = \omega(u, v)$ dla każdego $u, v \in V$.

Lemat 14. Niech (V, ω) będzie symplektyczną przestrzenią wektorową. Zbiór przekształceń symplektycznych z operacją składania tworzy grupę $\text{Sp}(V, \omega, 2n)$.

Dowód. Dla dowodu tezy lematu wystarczy pokazać, że zbiór przekształceń symplektycznych jest zamknięty ze względu na składanie i branie elementu odwrotnego. Rzeczywiście ustalmy dowolne dwa przekształcenia symplektyczne A_1 oraz A_2 . Wówczas dla przekształcenia $A_3 = A_2 \circ A_1$ z symplektyczności odwzorowań A_1 i A_2 dla dowolnych $u, v \in V$ otrzymujemy, że $\omega(A_3u, A_3v) = \omega(A_2(A_1u), A_2(A_1v)) = \omega(A_1u, A_1v) = \omega(u, v)$. Ponadto $\omega(A_1^{-1}u, A_1^{-1}v) = \omega(A_1(A_1^{-1}u), A_1(A_1^{-1}v)) = \omega(u, v)$, co dowodzi tezy lematu. \square

PRZYKŁAD 10.1 Przestrzeń \mathbb{R}^{2n} jest symplektyczną przestrzenią wektorową ze standardową formą symplektyczną ω zdefiniowaną dla dowolnych $x, y \in \mathbb{R}^{2n}$ wzorem

$$\omega(x, y) = x^T J y = \sum_{i,k=1}^{2n} x_i J_{ik} y_k.$$

Rzeczywiście forma ω jest dwuliniowa a ponadto $\omega(x, y) = x^T J y = (y^T J^T x)^T = -(y^T J x)^T = -(y^T J x) = \omega(y, x)$. Jeśli $\omega(x, y) = 0$ dla każdego $y \in \mathbb{R}^{2n}$, to podstawiając kolejno wektory bazy standardowej za y otrzymujemy, że $x = 0$. Co więcej, dla każdego $A \in \text{Sp}(2n, \mathbb{R})$ mamy, że $\omega(Ax, Ay) = x^T A^T J A y = x^T J y = \omega(x, y)$.

Twierdzenie 3. Niech ω będzie antysymetryczną i niezdegenerowaną formą dwuliniową na $k = 2n$ -wymiarowej przestrzeni wektorowej V nad \mathbb{R} . Wówczas istnieje baza $(e_1, \dots, e_k, f_1, \dots, f_k)$ (baza symplektyczna) taka, że $\omega(e_i, f_j) = \delta_{i,j}$ oraz $\omega(e_i, e_j) = \omega(f_i, f_j) = 0$.

Dowód. Przeprowadzimy dowód indukcyjny ze względu na k . Ze względu na niezdegenerowanie formy ω istnieją $u_1, v_1 \in V$ takie, że $\omega(u_1, v_1) = 1$ co implikuje, że podprzestrzeń $V_1 = \text{span}(u_1, v_1)$ jest symplektyczną przestrzenią wektorową. Zdefiniujmy podprzestrzeń liniową

$$W = \{v \in V : \forall w \in V_1 \omega(v, w) = 0\}$$

oraz zauważmy, że przestrzeń (W, ω) jest $2n - 2$ -wymiarową symplektyczną przestrzenią wektorową. Na podstawie indukcji istnieje baza symplektyczna $u_2, \dots, u_n, v_2, \dots, v_n$ przestrzeni W . Stąd $u_1, u_2, \dots, u_n, v_1, v_2, \dots, v_n$ jest bazą symplektyczną przestrzeni V . Ponadto $\omega = \sum_{k=1}^n du_k \wedge dv_k$ \square

PRZYKŁAD 10.2 Rozważmy funkcję Hamiltona H postaci $H(q, p) = K(p)$. Wówczas równania ruchu Hamiltona mają postać:

$$\begin{cases} \dot{q}_k = \frac{\partial H}{\partial p_k} = \frac{\partial K}{\partial p_k} = \omega_k(p) \\ \dot{p}_k = -\frac{\partial H}{\partial q_k} = 0 \end{cases},$$

gdzie $k = 1, \dots, n$. W konsekwencji rozwiązania powyższego układu są następujące:

$$p_k(t) = p_k^0, \quad q_k(t) = \omega_k(p^0)t + q_k^0, \quad 1 \leq k \leq n.$$

W rozwiązywaniu równań Hamiltona istotny jest wybór odpowiednich współrzędnych kanonicznych, w których układ równań przyjmuje prostszą postać.

PRZYKŁAD 10.3 Rozważmy równanie $\ddot{x} = -\omega^2 x$ oscylatora harmonicznego, gdzie $\omega^2 = k/m$ z hamiltonianem

$$H(x, p) = \frac{1}{2}(p^2 + \omega^2 x^2).$$

Równania Hamiltona przyjmują następującą postać:

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial p} = p \\ \dot{p} = -\frac{\partial H}{\partial x} = -\omega^2 x \end{cases}$$

Bez zmniejszenia ogólności możemy założyć, że $\omega = 1$. Szukamy przekształcenia kanonicznego, które pozwoli uzyskać zagadanie prostsze w rozwiązaniu. Wprowadźmy nowe współrzędne (φ, I) takie, że

$$p = f(I) \cos \varphi, \quad x = g(I) \sin \varphi,$$

a ponadto chcemy aby warunek kanoniczności był spełniony, tzn. aby zachodziło

$$dx \wedge dp = d\varphi \wedge dI.$$

Mamy

$$dx = g \cos \varphi d\varphi + g' \sin \varphi dI, \quad dp = -f \sin \varphi d\varphi + f' \cos \varphi dI,$$

gdzie u' oznacza pochodną u po I . Biorąc powyższe wzory otrzymujemy

$$\begin{aligned} dx \wedge dp &= g \cos \varphi f' \cos \varphi d\varphi \wedge dI - g' \sin \varphi f \sin \varphi dI \wedge d\varphi = \\ &= (gf' \cos^2 \varphi + g' f \sin^2 \varphi) d\varphi \wedge dI. \end{aligned}$$

Załóżmy, że $g = f$. Wówczas nasze przekształcenie będzie kanoniczne ma postać $f f' = 1$. Daje nam to rozwiązanie postaci $g(I) = f(I) = \sqrt{2I}$, co implikuje, że nowy Hamiltonian postaci

$$K(\varphi, I) = H(x(\varphi, I), p(\varphi, I)) = I.$$

Zatem równania Hamiltona w nowych zmiennych mają postać

$$\begin{cases} \dot{\varphi} = \frac{\partial K}{\partial I} = 1, \\ \dot{I} = -\frac{\partial K}{\partial \varphi} = 0. \end{cases}$$

Ich rozwiązania to $I(t) = I_0$ oraz $\varphi(t) = t + \varphi_0$.

PRZYKŁAD 10.4 Rozważmy ruch punktu materialnego w polu dwóch nieruchomych centrów grawitacyjnych. W punkcie $(-c, 0)$ umieszczamy masę m_1 , a w punkcie $(c, 0)$ masę m_2 , $c > 0$. Funkcja Hamiltona dla tego układu ma postać:

$$H(x, y, p_x, p_y) = \frac{1}{2}(p_x^2 + p_y^2) - \frac{Gm_1m_2}{\sqrt{(x-c)^2 + y^2}} + \frac{Gm_1m_2}{\sqrt{(x+c)^2 + y^2}}.$$

Aby móc rozwiązać równania ruchu należy wprowadzić współrzędne eliptyczne

$$\begin{cases} u = \sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} \\ v = \sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2}. \end{cases}$$

11. Nawiasy Poissona

Wiele własności układów hamiltonowskich wyraża się w terminach nawiasów Poissona. Niech $U \subset \mathbb{R}^{2n}$ będzie zbiorem otwartym oraz

$$\mathcal{F}(U) = \{f : U \rightarrow \mathbb{R} : f \in C^\infty(U, \mathbb{R})\}$$

Definicja 12. Niech $f, g \in \mathcal{F}(U)$. Nawias Poissona funkcji f i g definiujemy następująco

$$\begin{aligned} \{f, g\} &= \langle \nabla f, J \nabla g \rangle = \langle \nabla_q f, \nabla_p g \rangle - \langle \nabla_p f, \nabla_q g \rangle = \\ &= \sum_{k=1}^n \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right). \end{aligned}$$

Lemat 15. Niech $f, g, h \in \mathcal{F}(U)$ oraz $\alpha, \beta \in \mathbb{R}$. Wówczas

1. $\{f, g\} : U \rightarrow \mathbb{R}$ jest funkcją gładką,
2. $\{\alpha f + \beta g, h\} = \alpha \{f, h\} + \beta \{g, h\}$,
3. $\{f, g\} = -\{g, f\}$,
4. spełniona jest tożsamość Jacobiego

$$\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0,$$

5. spełniona jest tożsamość Leibniza

$$\{f, gh\} = g\{f, h\} + \{f, g\}h$$

6.

$$\frac{\partial}{\partial q_k}\{f, g\} = \left\{ \frac{\partial f}{\partial q_k}, g \right\} + \left\{ f, \frac{\partial g}{\partial q_k} \right\}.$$

Dowód. Ustalmy $f, g, h, \in \mathcal{F}(U)$. Dla dowodu punktu 5 zauważmy, że

$$\begin{aligned} \{f, gh\} &= \langle \nabla f, J\nabla(gh) \rangle = \langle \nabla f, J(h\nabla g + g\nabla h) \rangle = \\ &= h\langle \nabla f, J\nabla g \rangle + g\langle \nabla f, J\nabla h \rangle = h\{f, g\} + g\{f, h\}. \end{aligned}$$

Udowodnimy teraz punkt 6:

$$\begin{aligned} \frac{\partial}{\partial q_k}\{f, g\} &= \sum_{k=1}^n \frac{\partial}{\partial q_k} \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial f}{\partial p_k} \frac{\partial g}{\partial q_k} \right) = \sum_{k=1}^n \frac{\partial^2 f}{\partial q_k^2} \frac{\partial g}{\partial p_k} + \frac{\partial f}{\partial q_k} \frac{\partial^2 g}{\partial q_k \partial p_k} \\ &\quad - \frac{\partial^2 f}{\partial q_k \partial p_k} \frac{\partial g}{\partial q_k} - \frac{\partial f}{\partial p_k} \frac{\partial^2 g}{\partial q_k^2} = \left\{ \frac{\partial f}{\partial q_k}, g \right\} + \left\{ f, \frac{\partial g}{\partial q_k} \right\}. \end{aligned}$$

Pokażemy punkt 4:

$$\{f, \{g, h\}\} = \left\{ f, \sum_{k=1}^n \frac{\partial g}{\partial q_k} \frac{\partial h}{\partial p_k} - \frac{\partial g}{\partial p_k} \frac{\partial h}{\partial q_k} \right\} = \quad (11.1)$$

$$= \sum_{k=1}^n \left\{ f, \frac{\partial g}{\partial q_k} \frac{\partial h}{\partial p_k} \right\} - \sum_{k=1}^n \left\{ f, \frac{\partial g}{\partial p_k} \frac{\partial h}{\partial q_k} \right\} = \quad (11.2)$$

$$= \sum_{k=1}^n \frac{\partial g}{\partial q_k} \left\{ f, \frac{\partial h}{\partial p_k} \right\} + \frac{\partial h}{\partial p_k} \left\{ f, \frac{\partial g}{\partial q_k} \right\} - \frac{\partial g}{\partial p_k} \left\{ f, \frac{\partial h}{\partial q_k} \right\} - \frac{\partial h}{\partial q_k} \left\{ f, \frac{\partial g}{\partial p_k} \right\} = \quad (11.3)$$

$$\begin{aligned} &= \sum_{k=1}^n \frac{\partial g}{\partial q_k} \left(\frac{\partial}{\partial p_k} \{f, h\} - \left\{ \frac{\partial f}{\partial p_k}, h \right\} \right) + \frac{\partial h}{\partial p_k} \left(\frac{\partial}{\partial q_k} \{f, g\} - \left\{ \frac{\partial f}{\partial q_k}, g \right\} \right) + \\ &+ \frac{\partial g}{\partial p_k} \left(-\frac{\partial}{\partial q_k} \{f, h\} + \left\{ \frac{\partial f}{\partial q_k}, h \right\} \right) + \frac{\partial h}{\partial q_k} \left(-\frac{\partial}{\partial p_k} \{f, g\} + \left\{ \frac{\partial f}{\partial p_k}, g \right\} \right), \end{aligned} \quad (11.4)$$

gdzie 11.2 wynika z punktów (2) i (3), 11.3 wynika z punktu (5) a 11.4 z (6).

Zauważmy, że

$$\sum_{k=1}^n \frac{\partial g}{\partial q_k} \frac{\partial}{\partial p_k} \{f, h\} - \frac{\partial g}{\partial p_k} \frac{\partial}{\partial q_k} \{f, h\} = -\{g, \{h, f\}\}$$

oraz

$$\sum_{k=1}^n \frac{\partial h}{\partial p_k} \frac{\partial}{\partial q_k} \{f, g\} - \frac{\partial h}{\partial q_k} \frac{\partial}{\partial p_k} \{f, g\} = -\{h, \{f, g\}\},$$

a zatem

$$\begin{aligned}
& \{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = \\
& = \sum_{k=1}^n -\frac{\partial g}{\partial q_k} \left\{ \frac{\partial f}{\partial p_k}, h \right\} - \frac{\partial h}{\partial p_k} \left\{ \frac{\partial f}{\partial q_k}, g \right\} + \frac{\partial g}{\partial p_k} \left\{ \frac{\partial f}{\partial q_k}, h \right\} + \frac{\partial h}{\partial q_k} \left\{ \frac{\partial f}{\partial p_k}, g \right\} = \\
& = \sum_{k=1}^n \sum_{j=1}^n -\frac{\partial g}{\partial q_k} \frac{\partial^2 f}{\partial q_j \partial p_k} \frac{\partial h}{\partial p_j} + \frac{\partial g}{\partial q_k} \frac{\partial^2 f}{\partial p_j \partial p_k} \frac{\partial h}{\partial q_j} \\
& - \frac{\partial h}{\partial p_k} \frac{\partial^2 f}{\partial q_j \partial q_k} \frac{\partial g}{\partial p_j} + \frac{\partial h}{\partial p_k} \frac{\partial^2 f}{\partial p_j \partial q_k} \frac{\partial g}{\partial q_j} \\
& + \frac{\partial g}{\partial p_k} \frac{\partial^2 f}{\partial q_j \partial q_k} \frac{\partial h}{\partial p_j} - \frac{\partial g}{\partial p_k} \frac{\partial^2 f}{\partial p_j \partial q_k} \frac{\partial h}{\partial q_j} \\
& + \frac{\partial h}{\partial q_k} \frac{\partial^2 f}{\partial q_j \partial p_k} \frac{\partial g}{\partial p_j} - \frac{\partial h}{\partial q_k} \frac{\partial^2 f}{\partial p_j \partial p_k} \frac{\partial g}{\partial q_j} = 0.
\end{aligned}$$

□

12. Twierdzenie Liouville'a

Definicja 13. Całką układu (8.4) nazywamy funkcję $f \in C^1(U, \mathbb{R})$, która jest stała na wszystkich rozwiązaniach układu (8.4), tzn. rozwiązanie układu (8.4) nie zmienia poziomu całki.

Mówimy, że funkcje $f_1, \dots, f_l \in C^1(U, \mathbb{R})$ są funkcjonalnie niezależnymi całkami pierwszymi układu (8.4), o ile

$$\text{rank} \begin{bmatrix} \frac{\partial f_1}{\partial q_1} & \dots & \frac{\partial f_1}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_l}{\partial q_1} & \dots & \frac{\partial f_l}{\partial p_n} \end{bmatrix} = l.$$

Zatem funkcja f jest całką układu (8.4) wtedy i tylko wtedy, gdy pochodna f w kierunku pola wektorowego $J\nabla H$ jest tożsamościowo równa zero. Geometrycznie oznacza to, że dowolna trajektoria, która w pewnej chwili t_0 znajduje się na poziomie

$$M_c = \{x \in \mathbb{R}^{2n} : f(x) = c\}$$

w punkcie x_0 pozostaje na tej poziomie w każdym czasie. Jeżeli układ posiada l funkcjonalnie niezależnych całek pierwszych, to problem rozwiązania układu (8.4) sprowadza się do rozwiązywania układu na $(2n - l)$ -wymiarowej podrozmiarowości

$$M_c = \{x \in \mathbb{R}^{2n} : f_1(x) = c_1, \dots, f_l(x) = c_l\}.$$

W przypadku, gdy układ (8.4) posiada $2n - 1$ funkcjonalnie niezależnych całek pierwszych, to nazywany jest supercałkowalnym, zaś jednowymiarowa rozmiarowość $f_1^{-1}(c_1) \cap \dots \cap f_{2n-1}^{-1}(c_{2n-1})$ jest zbiorem niezmienniczym dla potoku generowanego przez układ (8.4).

Definicja 14. Układ hamiltonowski (8.4) nazywamy zupełnie całkownym, o ile istnieje n funkcjonalnie niezależnych całek układu $f_1, \dots, f_n \in C^1(U, \mathbb{R})$ będących w involucji, tj. $\{f_i, f_j\} = 0$ dla $i, j = 1, \dots, n$.

Twierdzenie 4. Ustalmy $f, g, h \in \mathcal{F}(U)$. Wówczas

- (1) funkcja f jest całką układu (8.4) wtedy i tylko wtedy, gdy $\{f, H\} = 0$,
- (2) hamiltonian H jest całką układu (8.4),
- (3) jeżeli funkcje f i g są całkami układu (8.4), to $\{f, g\}$ jest całką tego układu.

Dowód. Udowodnimy punkt (1). Niech $z : J_{(t_0, z_0)} \rightarrow U$ będzie rozwiązaniem układu (8.4) z warunkiem początkowym $z(t_0) = z_0$. Zauważmy, że

$$\frac{d}{dt}f(z(t)) = \langle \nabla f(z(t)), \dot{z}(t) \rangle = \langle \nabla f(z(t)), JH(z(t)) \rangle = \{f, H\}(z(t)).$$

Punkt (2) wynika z punktu (1) oraz tezy (7) lematu 15.

Dla dowodu punktu (3) założmy, że $\{f, H\} = \{g, H\} = 0$. Wówczas na podstawie tożsamości Jacobiego 15.(4)

$$\{H, \{f, g\}\} + \{g, \{H, f\}\} + \{f, \{g, H\}\} = 0,$$

co implikuje, że $\{H, \{f, g\}\} = 0$. □

Sformułujemy dalej fundamentalne twierdzenie Liouville'a.

Twierdzenie 5. (Liouville) Niech $f_1, \dots, f_n \in \mathcal{F}(U)$ będą funkcjonalnie niezależnymi całkami układu (8.4) w involucji oraz niech $M_c = f_1^{-1}(c_1) \cap \dots \cap f_{2n-1}^{-1}(c_{2n-1})$ będzie zbiorem zwartym i spójnym. Wówczas M_c jest dyfeomorficzny z n -wymiarowym torusem. Co więcej, na pewnym otoczeniu M_c istnieje kanoniczna zamiana zmiennych $(q, p) \rightarrow (I, \varphi)$ taka, że $\{I_i, I_j\} = \{\varphi_i, \varphi_j\} = 0$ oraz $\{\varphi_i, I_j\} = \delta_{i,j}$ dla $i, j = 1, \dots, n$ i zbiory poziomicowe postaci $I_1^{-1}(b_1) \cap \dots \cap I_n^{-1}(b_n)$ są torusami, a ponadto równania Hamiltona mają postać

$$\begin{cases} \dot{\varphi}_i &= \frac{\partial K}{\partial I_i} \\ \dot{I}_i &= 0 \end{cases}, i = 1, \dots, n,$$

gdzie $K = K(I)$.

13. Uogólnienie formalizmu Hamiltona

Definicja 15. Parę $(\mathbb{R}^k, \{\cdot, \cdot\})$ nazywamy k -wymiarową rozmaitością Poissona, o ile przekształcenie $\{\cdot, \cdot\} : \mathcal{F}(\mathbb{R}^k) \times \mathcal{F}(\mathbb{R}^k) \rightarrow \mathcal{F}(\mathbb{R}^k)$ jest dwuliniowe oraz spełnia następujące warunki:

- (1) $\{f, g\} = -\{g, f\}$ dla $f, g \in \mathcal{F}(\mathbb{R}^k)$,
- (2) $\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0$ dla $f, g, h \in \mathcal{F}(\mathbb{R}^k)$ (tożsamość Jacobiego),
- (3) $\{f, gh\} = g\{f, h\} + \{f, g\}h$ dla $f, g, h \in \mathcal{F}(\mathbb{R}^k)$ (własność Leibnitza).

PRZYKŁAD 13.1 (1) Para $(\mathbb{R}^{2n}, \{\cdot, \cdot\})$ ze standardowym nawiasem Poissona jest przykładem rozmaitości Poissona.

(2) Rozważmy przestrzeń \mathbb{R}^3 oraz przekształcenie zadane wzorem

$$\{f, g\}(x) = \left\langle x, \frac{\partial f}{\partial x} \times \frac{\partial g}{\partial x} \right\rangle$$

dla $f, g \in \mathcal{F}(\mathbb{R}^3), x \in \mathbb{R}^3$ (\times oznacza iloczyn wektorowy). Liniowość i antysymetryczność iloczynu skalarznego implikuje liniowość i asymetryczność przekształcenia $\{\cdot, \cdot\}$. Ponadto dla dowolnych $f, g, h \in \mathcal{F}(\mathbb{R}^3)$ i $x \in \mathbb{R}^3$ otrzymujemy

$$\{f, gh\}(x) = \left\langle x, \frac{\partial f}{\partial x} \times \left(\frac{\partial g}{\partial x} h + g \frac{\partial h}{\partial x} \right) \right\rangle = (\{f, g\}h + g\{f, h\})(x).$$

Wykonując odpowiednie rachunki sprawdzamy, że tożsamość Jacobiego również jest spełniona. Zatem para $(\mathbb{R}^3, \{\cdot, \cdot\})$ stanowi rozmaitość Poissona.

Definicja 16. Niech $f, g \in \mathcal{F}(\mathbb{R}^k)$ oraz $(\mathbb{R}^k, \{\cdot, \cdot\})$ będzie rozmaitością Poissona. Mówimy, że funkcje f i g są w inwolucji, o ile $\{f, g\} = 0$. Funkcję f , która jest w inwolucji z każdą funkcją z $\mathcal{F}(\mathbb{R}^k)$ nazywamy funkcją Casimira.

PRZYKŁAD 13.2 (1) Dla standardowej rozmaitości Poissona $(\mathbb{R}^{2n}, \{\cdot, \cdot\})$ funkcje Casimira są funkcje stałe.

(2) Funkcje Casimira stanowią funkcje radialne dla rozmaitości Poissona $(\mathbb{R}^3, \{\cdot, \cdot\})$ z przykładu 13.1. Istotnie, dla dowolnej funkcji radialnej $f(r) = f(\|x\|^2)$ mamy, że

$$\{f, g\}(x) = \left\langle x, 2 \frac{\partial f}{\partial r} x \times \frac{\partial g}{\partial x} \right\rangle = 2 \frac{\partial f}{\partial r} \left\langle x \times x, \frac{\partial g}{\partial x} \right\rangle = 0$$

dla dowolnego $g \in \mathcal{F}(\mathbb{R}^3)$, z własności iloczynu wektorowego $\langle a, b \times c \rangle = \langle a \times b, c \rangle$.

Definicja 17. Ustalmy $H \in \mathcal{F}(\mathbb{R}^k)$ na rozmaitości Poissona $(\mathbb{R}^k, \{\cdot, \cdot\})$. Równania postaci $\dot{x}_i = \{x_i, H\}$, $i = 1, \dots, k$ nazywamy układem równań Hamiltona.

PRZYKŁAD 13.3 (1) Dla standardowej rozmaitości Poissona mamy układ Hamiltona standardowej postaci: $\dot{q} = \frac{\partial H}{\partial p}$, $\dot{p} = -\frac{\partial H}{\partial q}$ dla $H \in \mathcal{F}(\mathbb{R}^{2n})$.

(2) Na rozmaitości Poissona jak w przykładzie 13.1 równania Hamiltona $\dot{x} = \{x, H\} = \frac{\partial H}{\partial x} \times x$ przyjmują następującą postać

$$\begin{cases} \dot{x}_1 = x_3 \frac{\partial H}{\partial x_2} - x_2 \frac{\partial H}{\partial x_3} \\ \dot{x}_2 = -x_3 \frac{\partial H}{\partial x_1} + x_1 \frac{\partial H}{\partial x_3} \\ \dot{x}_3 = x_2 \frac{\partial H}{\partial x_1} - x_1 \frac{\partial H}{\partial x_2} \end{cases}$$

dla $H \in \mathcal{F}(\mathbb{R}^3)$.

Twierdzenie 6. Ustalmy $f, g, H \in \mathcal{F}(U)$. Wówczas

1. Funkcja f jest całką układu $\dot{x} = \{x, H\}$, wtedy i tylko wtedy, gdy $\{f, H\} = 0$,

2. Hamiltonian H jest całką układu $\dot{x} = \{x, H\}$,
3. Dowolna funkcja Casimira f jest całką układu $\dot{x} = \{x, H\}$,
4. jeżeli funkcje f i g są całkami układu $\dot{x} = \{x, H\}$, to $\{f, g\}$ jest całką tego układu.

Twierdzenie 7. Ustalmy $H \in \mathcal{F}(\mathbb{R}^k)$ na rozamitości Poissona $(\mathbb{R}^k, \{\cdot, \cdot\})$. Wówczas istnieje zamiana zmiennych

$$x = f(y), \quad y = (q_1, \dots, q_r, p_1, \dots, p_r, c_1, \dots, c_l),$$

gdzie $k = 2r + l$ taka, że równania Hamiltona przyjmują następującą postać:

$$\begin{cases} \dot{q}_i = \frac{\partial K}{\partial p_i} \\ \dot{p}_i = -\frac{\partial K}{\partial q_i} \\ \dot{c}_j = 0 \end{cases}, \quad (13.1)$$

dla $i = 1, \dots, r, j = 1, \dots, l$; gdzie l jest liczbą funkcjonalnie niezależnych funkcji Casimira.

14. Stabilność Lapunowa

Niech dany będzie układ nieautonomiczny równań różniczkowych postaci

$$\dot{x}(t) = v(t, x(t)), \quad (14.1)$$

gdzie $U \subset \mathbb{R}^n$ i $J \subset \mathbb{R}$ są zbiorami otwartymi oraz $v \in C^1(U \times J, \mathbb{R}^n)$. Ponadto niech $x_0(t)$ będzie rozwiązaniem układu (14.1).

Definicja 18. Mówimy, że $x_0(t)$ jest stabilne w sensie Lapunowa dla $t \rightarrow +\infty$, jeśli dla każdego $\varepsilon > 0$ istnieje $t_0 \geq 0$ oraz $\delta > 0$ takie, że każde rozwiązanie $x(t)$ układu (14.1) takie, że $\|x(t_0) - x_0(t_0)\| < \delta$ spełnia dla $t > t_0$ warunek $\|x(t) - x_0(t)\| < \varepsilon$.

Jeśli dodatkowo $\lim_{t \rightarrow +\infty} \|x(t) - x_0(t)\| = 0$, to mówimy że $x_0(t)$ jest asymptotycznie stabilne w sesie Lapunowa.

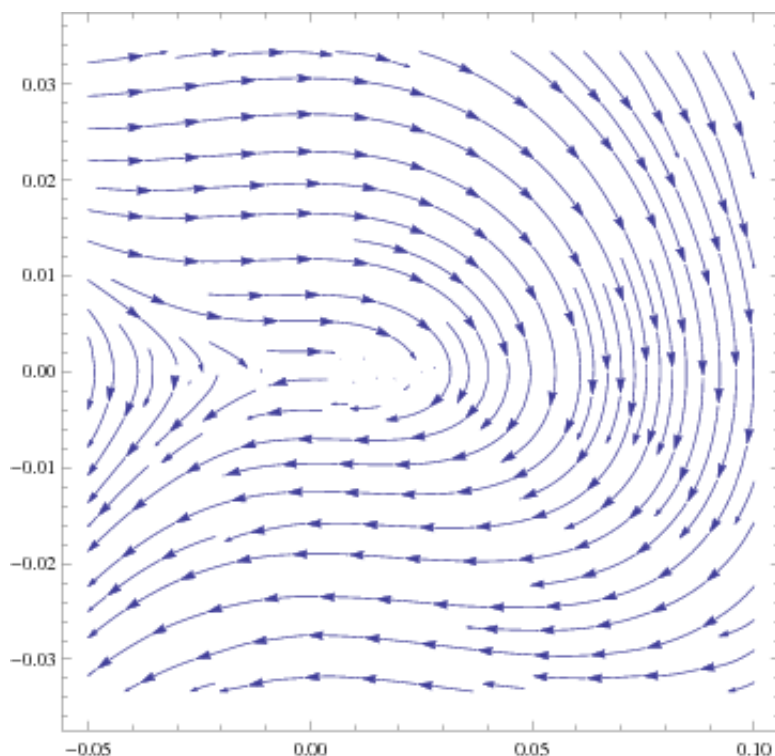
PRZYKŁAD 14.1 Rozważmy układ równań postaci

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -2x_1^2 - 4x_1x_2 - x_2(x_1^2 + x_2^2) \end{cases} \quad (14.2)$$

i zauważmy, że warunek $\lim_{t \rightarrow +\infty} \|x(t) - x_0(t)\| = 0$ z definicji asymptotycznej stabilności jest spełniony, ale rozwiązanie $x_0(t) \equiv 0$ nie jest stabilne w sensie Lapunowa.

Niech $x_0(t)$ będzie rozwiązaniem układu (14.1) oraz niech $y(t) = x(t) - x_0(t)$. Wówczas $\dot{x}(t) = \dot{y}(t) + \dot{x}_0(t)$. Stąd $f(t, x(t)) = \dot{y}(t) + f(t, x_0(t))$. Ostatecznie otrzymujemy

$$\dot{y}(t) = f(t, y(t) + x_0(t)) - f(t, x_0(t)). \quad (14.3)$$



Rysunek 3. Portret fazowy układu (14.2)

Reasumując, $x_0(t)$ jest rozwiązaniem układu (14.2) wtedy i tylko wtedy, gdy $y(t) \equiv 0$ jest rozwiązaniem układu (14.3). Innymi słowy, badanie stabilności i asymptotycznej stabilności rozwiązań równania różniczkowego sprowadza się do badania stabilności i asymptotycznej stabilności rozwiązania zerowego.

Rozważmy teraz układ autonomiczny równań różniczkowych postaci

$$\dot{x}(t) = v(x(t)), \quad (14.4)$$

gdzie $U \subset \mathbb{R}^n$ jest otoczeniem $0 \in \mathbb{R}^n$ oraz $v \in C^1(U \times J, \mathbb{R}^n)$ spełnia warunek $v(0) = 0$.

Definicja 19. Funkcją Lapunowa dla układu (14.4) nazywamy funkcję $L \in C^1(U, \mathbb{R})$ spełniającą następujące warunki:

- (1) $L(x) \geq 0$ dla każdego $x \in U$,
- (2) $L(x) = 0 \iff x = 0$,
- (3) jeżeli $x(t)$ jest rozwiązaniem układu (14.4), to funkcja $L(x(t))$ jest nierosnącą funkcją zmiennej t .

Zauważmy, że skoro $L(x(t))$ jest nierosnącą funkcją klasy C^1 , to zachodzi nierówność $\frac{d}{dt}L(x(t)) \leq 0$. Stąd

$$\begin{aligned} \frac{d}{dt}L(x(t)) &= \sum_{i=1}^n \frac{\partial L}{\partial x_i}(x(t)) \dot{x}_i(t) = \sum_{i=1}^n \frac{\partial L}{\partial x_i}(x(t)) v_i(x(t)) = \\ &= \langle \nabla L(x(t)), v(x(t)) \rangle \leq 0, \end{aligned}$$

dlatego zastąpimy warunek (3) z definicji 19 przez silniejszy warunek

$$\langle \nabla L(x(t)), v(x(t)) \rangle \leq 0$$

dla każdego $x \in U$, do sprawdzenia którego nie potrzebujemy znać rozwiązań układu (14.4). Wprowadzamy następujące oznaczenia: symbolem $B(x, r)$ (odpowiednio $D(x, r)$) oznaczamy kulę otwartą (domkniętą), o środku w punkcie x i promieniu r a przez $S(x, r)$ brzeg tej kuli.

Twierdzenie 8. (Lapunow) *Jeżeli dla układu (14.4) istnieje funkcja Lapunowa, to rozwiązanie $x_0 \equiv 0$ jest stabilne w sensie Lapunowa. Jeżeli ponadto zachodzi nierówność $\langle \nabla L(x(t)), v(x(t)) \rangle < 0$, to rozwiązanie $x_0 \equiv 0$ jest asymptotycznie stabilne w sensie Lapunowa.*

Dowód. Udowodnimy najpierw pierwszą część tezy. Ustalmy dowolne $\varepsilon > 0$ oraz niech $\delta = \min_{x \in S(0, \varepsilon)} L(x)$. Z warunków (1) oraz (2) definicji 19 i zwartości zbioru $S(0, \varepsilon)$ mamy, że $\delta > 0$. Niech

$$W = \{x \in D(0, \varepsilon) : L(x) < \delta\} = D(0, \varepsilon) \cap L^{-1}([0, \delta)).$$

Skoro $L(0) = 0$, to $W \neq \emptyset$. Jeżeli $x_1 \in W$, to z warunku (3) definicji 19 wynika, że rozwiązanie zagadnienia

$$\begin{cases} \dot{x} &= v(x(t)) \\ x(0) &= x_1 \end{cases}$$

nigdy nie osiągnie zbioru $S(0, \varepsilon)$, ponieważ dla każdego $t \in (0, t^+(x_0))$ mamy $L(x(t)) \leq L(x(0)) = L(x_1) < \delta$, co dowodzi, że rozwiązanie $x_0 \equiv 0$ jest stabilne w sensie Lapunowa. Dla dowodu drugiej części tezy wystarczy pokazać, że $\lim_{t \rightarrow +\infty} L(x(t)) = 0$. Przypuśćmy, że to jest nieprawdą, tzn. istnieje rozwiązanie $y(t)$ równania $\dot{x}(t) = v(x(t))$ takie, że $y(t_0) \in W$ oraz $\lim_{t \rightarrow +\infty} L(x(t)) \neq 0$. Funkcja $t \mapsto L(y(t))$ jest nierosnąca a zatem istnieje $\alpha > 0$ takie, że $L(y(t)) > \alpha > 0$ dla $t > t_0$. Ponadto istnieje $\varepsilon_1 > 0$ takie, że $\|y(t)\| > \varepsilon_1 > 0$ dla $t > t_0$, ponieważ $L(0) = 0$. Z warunku $\langle \nabla L(x(t)), v(x(t)) \rangle < 0$ oraz ze zwartości zbioru $D(0, \varepsilon) \setminus B(0, \varepsilon_1)$ wynika, że istnieje $m > 0$ takie, że $\frac{d}{dt} L(y(t)) \leq -m < 0$ dla $t > t_0$. Całkując powyższą nierówność w granicach od t_0 do t otrzymujemy, że

$$L(y(t)) - L(y(t_0)) \leq -m(t - t_0),$$

tj.

$$L(y(t)) \leq -m(t - t_0) + L(y(t_0)) \rightarrow -\infty$$

dla $t \rightarrow +\infty$, co przeczy warunkowi (2) definicji 19. \square

Niech $A \in M(n, \mathbb{R})$. Wprowadźmy następujące oznaczenia:

$$\begin{aligned} \sigma(A) &:= \{\lambda \in \mathbb{C} : \det(A - \lambda \text{Id}) = 0\}, \\ \sigma_-(A) &:= \{\lambda \in \lambda(A) : \text{Re}(\lambda) < 0\}, \\ \sigma_+(A) &:= \{\lambda \in \lambda(A) : \text{Re}(\lambda) > 0\}, \\ \sigma_0(A) &:= \{\lambda \in \lambda(A) : \text{Re}(\lambda) = 0\}. \end{aligned}$$

Dodatkowo symbolem $Gl(n, \mathbb{R})$ oznaczamy grupę odwracalnych macierzy rzeczywistych wymiaru $n \times n$.

Twierdzenie 9. Niech $A \in Gl(n, \mathbb{R})$ będzie takie, że $\sigma(A) = \sigma_-(A)$. Wówczas rozwiązanie $x_0 \equiv 0$ układu równań postaci $\dot{x}(t) = Ax(t)$ jest asymptotycznie stabilne w sensie Lapunowa.

Dowód. Z twierdzenia Jordana mamy, że $A = PJ(A)P^{-1}$, gdzie $J(A)$ jest macierzą Jordana macierzy A . Ustalmy $x_1 \in \mathbb{R}^n$ i rozważmy zagadnienie początkowe następującej postaci

$$\begin{cases} \dot{x} &= Ax(t) \\ x(0) &= x_1 \end{cases}.$$

Wówczas rozwiązanie zagadnienia początkowego ma postać

$$x(t) = e^{At}x_1 = Pe^{J(A)t}P^{-1}x_1,$$

gdzie $J(A) = \text{diag}(B_1, \dots, B_r)$ jest macierzą Jordana macierzy A . Ustalmy $\mu > 0$ takie, że dla każdego $\lambda \in \sigma(A)$ mamy $\text{Re}(\lambda) < -\mu$. Oznaczmy przez $\widetilde{J(A)}$ macierz otrzymaną z $J(A)$ poprzez zastąpienie wszystkich elementów na głównej przekątnej przez $-\mu$. Wówczas otrzymujemy

$$\begin{aligned} \|x(t)\| &= \|e^{At}x_1\| = \|Pe^{J(A)t}P^{-1}x_1\| \leq \|P\| \|e^{J(A)t}P^{-1}x_1\| \leq \\ &\leq \|P\| \|e^{\widetilde{J(A)}t}P^{-1}x_1\| = e^{-\mu t} \|P\| \|e^{t(\widetilde{J(A)} + \mu Id)}P^{-1}x_1\| \leq \\ &\leq \|P\| \|P^{-1}\| \|e^{-\mu t}e^{t(\widetilde{J(A)} + \mu Id)}\| \|x_1\|. \end{aligned}$$

Elementami macierzy $e^{t(\widetilde{J(A)} + \mu Id)}$ są kombinacje liniowe iloczynów wielomianów oraz funkcji $\sin \beta t$ oraz $\cos \beta t$, dlatego

$$\lim_{t \rightarrow +\infty} e^{-\mu t} e^{t(\widetilde{J(A)} + \mu Id)} = 0 \in M(n, \mathbb{R}). \quad (14.5)$$

A w konsekwencji $\lim_{t \rightarrow +\infty} x(t) = 0$. Ponadto z warunku 14.5 otrzymujemy, że dla każdego $\varepsilon > 0$ istnieje $t_0 > 0$ takie, że dla każdego $t \geq t_0$ zachodzi $\|P\| \|P^{-1}\| \|e^{-\mu t} e^{t(\widetilde{J(A)} + \mu Id)}\| \|x_1\| < \varepsilon$. Reasumując otrzymaliśmy, że dla każdego $\varepsilon > 0$ istnieje $t_0 > 0$ oraz $\delta = 1$ takie, że dla każdego $t \geq t_0$ oraz $x_1 \in \mathbb{R}^n$ takich, że $\|x_1\| < 1$ mamy

$$\begin{aligned} \|x(t)\| &\leq \|P\| \|P^{-1}\| \|e^{-\mu t} e^{t(\widetilde{J(A)} + \mu Id)}\| \|x_1\| < \\ &< \|P\| \|P^{-1}\| \|e^{-\mu t} e^{t(\widetilde{J(A)} + \mu Id)}\| < \varepsilon, \end{aligned}$$

co kończy dowód twierdzenia. \square

Twierdzenie 10. Niech $A \in M(n, \mathbb{R})$ będzie takie, że $\sigma_+(A) \neq \emptyset$. Wówczas rozwiązanie $x_0 \equiv 0$ układu równań $\dot{x}(t) = Ax(t)$ nie jest stabilne w sensie Lapunowa.

Twierdzenie 11. Niech $A \in M(n, \mathbb{R})$. Wówczas, jeśli rozwiązanie $x_0 \equiv 0$ układu równań $\dot{x}(t) = Ax(t)$ jest stabilne w sensie Lapunowa, to $\sigma(A) = \sigma_-(A) \cup \sigma_0(A)$. Jeśli albo $\sigma(A) = \sigma_-(A)$ albo $\sigma(A) = \sigma_-(A) \cup \sigma_0(A)$ i wartości własne o zerowej części rzeczywistej mają tylko jednowymiarowe klatki Jordana, to rozwiązanie $x_0 \equiv 0$ jest stabilne w sensie Lapunowa.

Twierdzenie 12. Niech $U \subset \mathbb{R}^n$ jest otoczeniem $0 \in \mathbb{R}^n$ oraz $v \in C^2(U, \mathbb{R}^n)$ takie, że $v(0) = 0$. Rozważmy układ postaci $\dot{x}(t) = Ax(t) + o(\|x\|^2)$ dla x z pewnego otoczenia $0 \in \mathbb{R}^n$, gdzie $A = \frac{\partial v}{\partial x}(0)$. Wówczas

- (1) jeżeli istnieje funkcja Lapunowa L dla układu $\dot{x}(t) = Ax(t)$, to rozwiązanie $x_0 \equiv 0$ jest stabilne w sensie Lapunowa. Jeżeli ponadto zachodzi nierówność $\langle \nabla L(x), x \rangle < 0$, to rozwiązanie $x_0 \equiv 0$ jest asymptotycznie stabilne w sensie Lapunowa,
- (2) jeżeli $\sigma(A) = \sigma_-(A)$, to rozwiązanie $x_0 \equiv 0$ jest asymptotycznie stabilne w sensie Lapunowa,
- (3) jeżeli $\sigma_+(A) \neq \emptyset$, to rozwiązanie $x_0 \equiv 0$ nie jest stabilne w sensie Lapunowa.

15. Stabilność położeń równowagi układów hamiltonowskich

Rozważmy układ równań postaci

$$\dot{x}(t) = v(x) = J \frac{\partial h}{\partial x}(x(t)), \quad (15.1)$$

gdzie $U \subset \mathbb{R}^{2n}$ jest otoczeniem $0 \in \mathbb{R}^{2n}$ oraz $h \in C^\infty(U, \mathbb{R})$. Położeniami równowagi układu są punkty krytyczne funkcji h . Załóżmy, że $\frac{\partial h}{\partial x}(0) = 0$. Będziemy zajmować się linearyzacją układu (15.1) w punkcie $x_0 = 0$ to znaczy

$$\dot{x}(t) = Ax(t), \quad \text{gdzie} \quad A = J \frac{\partial^2 h}{\partial x^2}(0) = JH. \quad (15.2)$$

Oczywiście macierz Hessego H jest symetryczna, tzn. $H = H^T$. Zakładając, że macierz H jest dodatnio określona otrzymujemy, że funkcja $L = h|_W : W \rightarrow \mathbb{R}$ jest funkcją Lapunowa (15.1), gdzie $W \subset \mathbb{R}^{2n}$ jest pewnym otoczeniem $0 \in \mathbb{R}^{2n}$, a stąd mamy stabilność rozwiązania $x_0 \equiv 0$ w sensie Lapunowa. Rzeczywiście, bez zmniejszenia ogólności możemy założyć, że $h(0) = 0$. Ponadto z dodatniej określoności H mamy lokalne minimum w punkcie $0 \in \mathbb{R}^{2n}$, tj. $L(x) \geq 0$ dla każdego $x \in W$, oraz

$$\langle \nabla h(x(t)), \dot{x}(t) \rangle = \langle \nabla h(x(t)), J \frac{\partial h}{\partial x}(x(t)) \rangle = 0.$$

Będziemy teraz badać stabilność rozwiązania $x_0 \equiv 0$ układu (15.2). w tym celu szukamy wartości własnych macierzy $A = JH$. Zauważmy, że

$$\begin{aligned} p(\lambda) &= \det(JH - \lambda \text{Id}_{2n}) = \det(JH - \lambda \text{Id}_{2n})^T = \det(-HJ - \lambda \text{Id}_{2n}) = \\ &= \det(J^2 HJ + \lambda J^2) = (\det J)^2 \det(JH + \lambda \text{Id}_{2n}) = p(-\lambda), \end{aligned}$$

ponieważ $J^T = -J$, $J^2 = -\text{Id}_{2n}$ oraz $\det(J^2) = (-1)^{2n} = 1$. Zatem $\lambda \in \sigma(A) \iff -\lambda \in \sigma(A)$, co więcej $\lambda \in \sigma(A) \iff \bar{\lambda} \in \sigma(A)$,

Twierdzenie 13. (Hartman-Grobman) Niech $U \subset \mathbb{R}^n$ będzie otoczeniem punktu $x_0 \in U$ oraz niech $v \in C^1(U, \mathbb{R}^n)$. Przez $\phi(\cdot, \cdot)$ oznaczamy lokalny układ dynamiczny indukowany przez układ równań $\dot{x}(t) = v(x(t))$. Załóżmy, że punkt $x_0 \in U$ jest hiperbolicznym punktem stacjonarnym układu $\dot{x}(t) = v(x(t))$ (tzn. $\sigma(Dv(x_0)) \cap i\mathbb{R} = \emptyset$). Wówczas równania $\dot{x}(t) = v(x(t))$ oraz $\dot{x}(t) = Dv(x_0)x(t)$ są topologicznie sprzężone w otoczeniu punktu $x_0 \in \mathbb{R}^n$, tzn. istnieje otoczenie $W \subset U$ punktu $x_0 \in U$ oraz homeomorfizm $F : W \rightarrow F(W)$ taki, że dla dowolnego $\xi \in W$ istnieje $J(\xi) \subset \mathbb{R}$ taki, że dla każdego $t \in J(\xi)$ zachodzi równość $F \circ \phi(t, \xi) = e^{Dv(x_0)t} F(\xi)$.

Na podstawie twierdzenia 13, gdy $0 \in \mathbb{R}^{2n}$ jest hiperbolicznym punktem stacjonarnym układu (15.1), to układy (15.1) oraz (15.2) mają tę samą własność stabilności rozwiązania $x_0 \equiv 0$. Natomiast jeżeli $\sigma(A) \cap i\mathbb{R} \neq \emptyset$, to o stabilności Lapunowa rozwiązania $x_0 \equiv 0$ decyduje nie tylko linearyzacja, lecz także człony wyższych rzędów w rozwinięciu Taylora funkcji $v(x)$. Reasumując, badanie stabilności Lapunowa rozwiązań układu (15.1) metodami Lapunowa nie jest proste.

Opiszemy teraz metodę badania stabilności rozwiązania $x_0 \equiv 0$ układu (15.1) teorią postaci normalnych pochodzącą od Poincaré, w przypadku gdy $0 \in \mathbb{R}^{2n}$ nie jest hiperbolicznym punktem stacjonarnym układu. Upraszczamy sytuację zakładając, że $\sigma(A) = \{\lambda_1, \dots, \lambda_n, -\lambda_1, \dots, -\lambda_n\}$, gdzie $\text{Re}(\lambda_k) = 0$ dla $k = 1, \dots, n$ oraz $\lambda_i \neq \lambda_j$ dla $i \neq j$, tj. $\lambda_k = +i\sigma_k$, $\lambda_{k+n} = -i\sigma_k$, $\sigma_k \neq 0$ dla $k = 1, \dots, n$ oraz $\sigma_i \neq \sigma_j$ dla $i \neq j$. Przeprowadzimy zamianę zmiennych $x = Sy$ taką, że linearyzacja układu (15.1) w nowych zmiennych ma postać $\dot{y}(t) = J\tilde{H}y(t)$, gdzie $\tilde{H} = \text{diag}(\sigma_1, \dots, \sigma_n, \sigma_1, \dots, \sigma_n)$ oraz S spełnia warunek

$$S^T J S = J. \quad (15.3)$$

Zatem $\dot{x} = S\dot{y}$ dalej $JHSy = SJ\tilde{H}y$ a stąd $JHS = SJ\tilde{H}$. Niech $S = BC$, gdzie $B, C \in \text{Gl}(2n, \mathbb{R})$. Wówczas $JHBC = BCJ\tilde{H}$ lub równoważnie $JHB = BCJ\tilde{H}C^{-1}$. Ustalmy

$$C = \begin{bmatrix} i\text{Id}_{2n} & \text{Id}_{2n} \\ -i\text{Id}_{2n} & \text{Id}_{2n} \end{bmatrix}.$$

Wtedy

$$CJ\tilde{H}C^{-1} = \begin{bmatrix} i\Sigma & \Theta_n \\ \Theta_n & -i\Sigma \end{bmatrix} = \text{diag}(i\sigma_1, \dots, i\sigma_n, -i\sigma_1, \dots, -i\sigma_n) = D.$$

Ustalmy teraz $B = [e_1, \dots, e_n, e_{n+1}, \dots, e_{2n}]$, gdzie

$$\begin{aligned} JHe_k &= i\sigma_k e_k \\ JHe_{n+k} &= -i\sigma_k e_{n+k} \end{aligned} \quad (15.4)$$

dla $k = 1, \dots, n$. Wtedy warunek (15.3) przyjmuje postać $J = C^T B^T J B C = C^T F C$, gdzie

$$F = [e_k^T J e_l] = \begin{bmatrix} f^1 & f^3 \\ -f^3 & f^2 \end{bmatrix},$$

tj.

$$\begin{cases} f^1 + f^2 & = \Theta_n \\ i(f^1 + 2f^3 - f^2) & = \text{Id}_n \\ i(f^1 + 2f^3 - f^2) & = \text{Id}_n \end{cases} \iff \begin{cases} f^1 = f^2 & = \Theta_n \\ f^3 & = -\frac{1}{2}i \text{Id}_n \end{cases}.$$

Stąd otrzymujemy, że $F = [e_k^T J e_l] = -\frac{1}{2}iJ$. Ponadto, jeśli wprowadzimy następujące oznaczenia: $e_k = r_k + is_k$ oraz $e_{n+k} = r_k - is_k$ dla $k = 1, \dots, n$, to warunek (15.4) jest równoważny następującemu układowi:

$$\begin{cases} JHr_k & = -\sigma_k s_k \\ JHs_k & = \sigma_k r_k \end{cases}$$

dla $k = 1, \dots, n$, a stąd

$$-\frac{1}{2}i = (r_k + is_k)^T J (r_k - is_k) = -ir_k^T J s_k + is_k^T J r_k,$$

ponieważ $r_k^T J r_k = s_k^T J s_k = 0$. Równoważnie otrzymujemy, że $1 = 4r_k^T J s_k$. Zatem stosując zamianę zmiennych S w układzie (15.1) otrzymujemy, że

$$\dot{y} = J\tilde{H}y + \sum_{m=3}^M \frac{1}{(m-1)!} S^{-1} J \frac{\partial^m h}{\partial x^m}(0) (Sy)^{m-1} + o(\|y\|^{M+1}), \quad (15.5)$$

gdzie hamiltonian stowarzyszony z powyższym układem równań ma postać

$$h(y) = \frac{1}{2} \sum_{k=1}^n \sigma_k (y_k^2 + y_{n+k}^2) + \sum_{m=3}^M h_m(y) + o(\|y\|^{M+1}). \quad (15.6)$$

W przypadku równań Hamiltona definicja rezonansu jest inna niż ta stosowana w teorii Poincaré.

Definicja 20. Wartości własnych $\sigma = (\sigma_1, \dots, \sigma_n, -\sigma_1, \dots, -\sigma_n)$ macierzy $A = J\tilde{H}$ nazywamy rezonansowymi, gdy $(m_1, \dots, m_n) \in \mathbb{Z}^n$ takie, że

$$\sum_{i=1}^n m_i \sigma_i = 0, \quad \sum_{i=1}^n |m_i| = k > 0 \quad (15.7)$$

Liczbę k nazywamy rzędem rezonansu.

Stosując kolejne zamiany zmiennych można dalej upraszczać funkcję Hamiltona. Cała procedura bardzo silnie zależy od tego czy występują rezonanse czy też nie. Jeżeli nie ma rezonansów do czwartego rzędu włącznie, to kanoniczną zamianą zmiennych możemy usunąć wszystkie wyrazy

trzeciego rzędu oraz bardzo uprościć wyrazy rzędu 4. W efekcie postać normalna funkcji Hamiltona będzie następująca

$$h(y) = \sum_{k=1}^n \sigma_k I_k + \sum_{k,l=1}^n a_{k,l} I_k I_l + \dots \quad (15.8)$$

gdzie

$$I_k = \frac{1}{2} (y_k^2 + y_{n+k}^2), \quad k = 1, \dots, n.$$

Taka właśnie postać normalna wykorzystywana jest do badania stabilności.

Literatura

- [1] V.I. Arnold, *Teoria równań różniczkowych*, Warszawa, PWN (1983)
- [2] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag (1978)
- [3] H. Goldstein, *Classical Mechanics*, Addison-Wesley (1980)
- [4] F. Scheck, *Mechanics: From Newton's Laws to Deterministic Chaos*, Springer Berlin Heidelberg (2010)