PODSTAWY MECHANIKI KLASYCZNEJ
wersja robocza

Andrzej Pękalski

Spis treści

Rozdział 1. Wstęp 5

Rozdział 2. Dynamika punktu materialnego 7
1. Zasady zachowania 11
2. Układ punktów materialnych 13
3. Więzy 14
4. Ogólne równanie dynamiki. 17
5. Równania Lagrange'a II rodzaju. 17
6. Prawa zachowania we współrzędnych uogólnionych 23

Rozdział 3. Zasada Hamiltona 27

Rozdział 4. Formalizm Hamiltona 31
1. Równania Hamiltona 31
2. RH z zasady Hamiltona 35
3. Nawiasy Poissona 35

Rozdział 5. Zagadnienie dwu ciał 39
1. Równania toru 39
2. Prawa Keplera 42

Rozdział 6. Szczególna teoria względności 45
1. Wydłużenie czasu i skrócenie lorentzowskie 46
2. Przekształcenie Lorentza 48
3. Wnioski z przekształcenia Lorentza 51
4. Przedziały między zdarzeniami i ich klasyfikacja 53
5. Czterowektory i równania ruchu 55

ROZDZIAŁ 1

Wstęp

Mechanika klasyczna, zwana także teoretyczną lub analityczną, to dział fizyki teoretycznej zajmujący się badaniem praw ruchu ciał materialnych. Jest to mechanika Newtonowska, której podstawy zostały stworzone przed powstaniem mechaniki kwantowej i teorii względności. Zagadnienia do których stosują się metody rozwinięte tutaj można określić jako „zwyczajne", tzn zagadnienia dotyczące ruchu ciał o niezbyt małych rozmiarach i poruszających się z prędkościami małymi w porównaniu z prędkością światła. Pojawiające się czasem różnice pomiędzy teorią i doświadczeniem również i w tym zakresie zjawisk, należy położyć na karb przyjęcia zbyt prostego modelu matematycznego - zaniedbanie tarcia, zastąpienie ciała sprężystego przez sztywne itd. Kiedy względne prędkości ciał stają się porównywalne z prędkością światła lub kiedy rozmiary ciał są rzędu stałych atomowych, wówczas mechanika klasyczna nie opisuje już zjawisk w sposób poprawny. W pierwszym przypadku należy ją zastąpić przez mechanikę relatywistyczną, w drugim przez mechanikę kwantową.

Aczkolwiek mechanika klasyczna znajduje dzisiaj ciągle jeszcze bezpośrednie zastosowanie, np w astronomii, to jednak główna jej wartość polega na tym, że jest ona podstawą całej fizyki teoretycznej. Wiele pojęć np mechaniki kwantowej określanych jest poprzez analogie z odpowiednimi pojęciami mechaniki klasycznej.

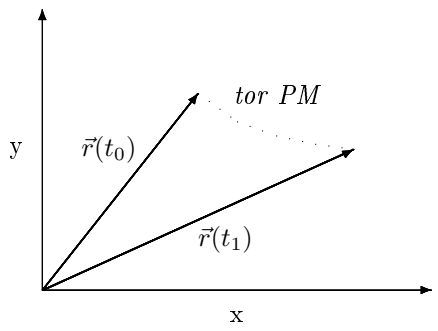
W mechanice klasycznej, tak jak w całej fizyce teoretycznej, operujemy modelami matematycznymi. Zamieniamy w nich istniejącą realną rzeczywistość poprzez pewien uproszczony model. Robimy tak głównie dlatego, że uwzględnienie wszystkich czynników wpływających na rzeczywistą sytuację nie jest ani możliwe, ani też na ogół potrzebne. Dlatego też musimy uwzględnić tylko niektóre z tych czynników. Takie które są, lub które uważamy za najistotniejsze. Rozwiązanie tak postawionego problemu powinno poprawnie opisywać najważniejsze cechy charakterystyczne danego procesu. Generalną zasadą jest aby początkowo konstruować możliwie prosty model i dopiero po sprawdzeniu jego ogólnej przydatności wprowadzać dodatkowe parametry umożliwiające opis bardziej szczegółowych własności. Dlatego też może istnieć kilka modeli teoretycznych opisujących tę samą rzeczywistość.

Pojęciem, którym będziemy się cały czas posługiwali jest pojęcie **punktu materialnego** (PM). Rozumiemy przez to ciało fizyczne o rozmiarach pomijalnie małych w porównaniu i innymi rozmiarami z jakimi mamy do czynienia w rozpatrywanym zagadnieniu. Będzie to więc punkt matematyczny, któremu została przypisana pewna masa. W tym kontekście PM może być np Ziemia, jeżeli rozpatrujemy jej ruch roczny wokół Słońca. Położenie PM w danym układzie współrzędnych, lub jak teraz będziemy mówić **układzie odniesienia**, podawać będziemy przy pomocy **promienia wodzącego** $\vec{r}$, o współrzędnych (x, y, z). Jest to wektor poprowadzony z początku układu do miejsca w którym znajduje się PM.

Aby określić ruch PM podajemy zależność promienia wodzącego od czasu $t : \vec{r} = \vec{r}(t)$. Krzywą geometryczną zakreślaną przez PM w czasie ruchu nazywamy **torem** PM lub jego **trajektorią**. Równanie

$$\vec{r} = \vec{r}(t)$$

jest parametrycznym równaniem toru, przy czym parametrem jest czas t. O funkcji $\vec{r}$ zakładamy, że jest ciągła i dwukrotnie różniczkowalna (klasy C^2). Założenia te usprawiedliwione są poprzez zgodność otrzymanych wniosków z doświadczeniem.



Rysunek 1. Wektory wodzące w dwu chwilach czasu.

Prędkością PM $\vec{v}$, nazywamy pochodną względem czasu promienia wodzącego

$$\vec{v} = \frac{d\vec{r}}{dt} \equiv \dot{\vec{r}}.$$

Przyspieszeniem PM $\vec{a}$, nazywamy wektor będący pochodną względem czasu prędkości

$$\vec{a} = \frac{d\vec{v}}{dt} \equiv \dot{\vec{v}} \equiv \ddot{\vec{r}}.$$

Kropka nad literą będzie zawsze oznaczała pochodną względem czasu.

ROZDZIAŁ 2

Dynamika punktu materialnego

Zajmować się teraz będziemy ruchem punktu materialnego (PM) pod wpływem zewnętrznej siły. Podstawą rozważań są prawa ruchu Newtona otrzymane jako uogólnienie danych doświadczalnych.

Pierwsze prawo Newtona

Istnieje układ odniesienia, w którym PM porusza się bez przyspieszenia (tzn jednostajnie i prostoliniowo), jeżeli nie działa nań z zewnątrz żadna siła, lub też działające siły równoważą się.

Układy, o których mowa w tym punkcie nazywamy układami *inercjalnymi*. Jeżeli układ U' porusza się względem układu inercjalnego U ruchem prostoliniowym i jednostajnym, wówczas i układ U' jest inercjalny. Słuszne jest również i twierdzenie odwrotne, tzn dwa układy inercjalne mogą poruszać się względem siebie tylko ruchem jednostajnym i prostoliniowym. W dalszym ciągu wykładu, o ile nie zostanie to osobno podkreślone, rozpatrywać będziemy tylko układy inercjalne.

Drugie prawo Newtona

Można je traktować jako definicję siły i masy:

$$\vec{F} = m\vec{a} = m\frac{d\vec{v}}{dt} = m\ddot{\vec{r}}. \tag{2.1}$$

Wybierając współczynnik proporcjonalności między siłą F i przyspieszeniem a, decydujemy się na pewien układ jednostek. Przyjęliśmy tutaj (co nie zawsze jest słuszne), że ten współczynnik nie zależy od czasu. Można więc drugie prawo Newtona zapisać w postaci:

$$\vec{F} = \frac{d(m\vec{v})}{dt} = \frac{d\vec{p}}{dt}, \tag{2.2}$$

gdzie wielkość $\vec{p} = m\vec{v}$ jest pędem PM. Przez $\vec{F}$ rozumiemy wypadkową siłę działającą na PM. Dla układów inercjalnych I prawo jest konsekwencją II. Jeśli bowiem wypadkowa siła $\vec{F} = 0$, wówczas $m\vec{a} = 0$, czyli jeśli $m \neq 0$ to $\vec{a} = 0$ i PM porusza się bez przyspieszenia.

Ruch PM dany jest przez wektorową zależność funkcyjną $\vec{r} = \vec{r}(t)$. Możemy ją znaleźć z II prawa Newtona, traktując $\vec{F}$ jako wielkość znaną z doświadczenia. Układ równań (2.2), lub we współrzędnych kartezjańskich

$$m\ddot{x} = F_x, \quad m\ddot{y} = F_y, \quad m\ddot{z} = F_z, \tag{2.3}$$

nazywamy **równaniami ruchu Newtona**. Mając daną funkcję $\vec{F} = (F_x, F_y, F_z)$, poprzez całkowanie równań (2.2) otrzymujemy zależność

$$\vec{r} = \vec{r}(t), \quad \text{lub} \quad x = x(t), \quad y = y(t), \quad z = z(t),$$

określającą ruch PM. Znalezienie takiej zależności jest równoznaczne z wyznaczeniem toru PM, jeżeli znane są odpowiednie warunki początkowe, i stanowi zasadniczy cel mechaniki klasycznej.

Trzecie prawo Newtona

Siły oddziaływania dwu PM są równe co do wielkości, mają przeciwny zwrot i działają wzdłuż prostej łączącej te punkty.

Prawo to nie jest uniwersalne i nie jest słuszne np. w przypadku sił elektromagnetycznych pomiędzy

poruszającymi się ładunkami.
Zasada niezależności sił, sformułowana przez Newtona w oparciu o dane doświadczalne, mówi że:
Siły działają na PM niezależnie, tzn dodanie nowej siły nie zmienia dotychczas działających.
Tak więc siły, a więc i przyspieszenia, możemy traktować jak wektory i dodawać do siebie geometrycznie.

Jeśli znana jest siła $\vec{F}$ działająca na PM, to korzystając z równań Newtona (2.2) lub (2.3) i całkując je powinniśmy otrzymać, przy zadanych warunkach początkowych, rozwiązanie zagadnienia, tzn tor PM. Dla przykładu znajdziemy tor PM pod działaniem siły ciężkości Ziemi. Załóżmy, że oś O_x skierowana jest pionowo w górę. Mamy wtedy

$$F_x = -mg, \quad F_y = 0, \quad F_z = 0, \quad g = const.$$

Równania Newtona mają postać

$$m\frac{d^2x}{dt^2} = -mg, \quad \frac{d^2y}{dt^2} = \frac{d^2z}{dt^2} = 0. \tag{2.4}$$

Jak widać, istotna jest tylko zależność x od t. Dlatego też dalej rozpatrzymy tylko pierwsze równanie. Wycałkujmy je od pewnej początkowej chwili czasu t_0 do chwili końcowej t. Otrzymamy

$$\frac{dx(t)}{dt} - \frac{dx(t_0)}{dt} = -\int_{t_0}^{t} g\,dt = -g(t - t_0).$$

Po powtórnym scałkowaniu dostaniemy

$$x(t) = \frac{dx(t_0)}{dt}(t - t_0) + x(t_0) - \frac{g}{2}(t - t_0)^2. \tag{2.5}$$

Otrzymaliśmy więc szukaną zależność $x = x(t)$ w postaci całki ogólnej równania Newtona. Aby jednak wyznaczyć do końca ruch PM (znaleźć rozwiązanie szczególne) należy zadać warunki początkowe. Równanie jakie otrzymaliśmy powinno opisywać dowolną sytuację w której na PM działa tylko siła grawitacji Ziemi. Zadanie warunków początkowych wybiera spośród wszystkich możliwych rozwiązań jedno, odpowiadające konkretnemu zagadnieniu. Niech w naszym przypadku będzie nim swobodne spadanie. Tak więc w chwili początkowej t_0 PM znajduje się w spoczynku na pewnej wysokości h. Rachubę czasu rozpocznijmy od chwili t_0 . Nasze warunki początkowe będą więc miały postać

$$x(t_0) = x(0) = h, \quad \frac{dx(t_0)}{dt} = v_0 = 0, \quad t_0 = 0.$$

Podstawiając te warunki do znalezionego rozwiązania (2.5) dostajemy

$$x(t) = h - \frac{gt^2}{2} \quad \text{oraz} \quad \frac{dx(t)}{dt} = v(t) = -gt.$$

Rozpatrzmy teraz rzut w górę. Oznacza to, że w chwili początkowej t_0 ($= 0$ dla prostoty) PM znajduje się w położeniu $x(0) = x_0$, które też bez straty ogólności możemy przyjąć za równe 0. Prędkość początkowa jest teraz różna od zera, tzn $dx(t_0)/dt = v_0 > 0$. Podstawiając te warunki do rozwiązania (2.5)dostajemy

$$x(t) = v_0 t - \frac{gt^2}{2}, \quad v(t) \equiv \frac{dx(t)}{dt} = v_0 - gt \tag{2.6}$$

Stąd już łatwo możemy znaleźć np maksymalna wysokość (x_m) jaką osiągnie PM. W punkcie tym prędkość punktu zmienia znak i jest równa zero, $v(t_m) = 0$, gdzie t_m to czas w jakim to zajdzie. Czas ten możemy wyznaczyć korzystając z drugiego równania w (2.6)

$$v(t_m) \;=\; 0 \;=\; v_0 - gt_m \quad \Rightarrow \quad t_m \;=\; \frac{v_0}{g}.$$

Wstawiając to do równania toru, dostajemy

$$x_m \equiv x(t_m) = v_0 t_m - \frac{g t_m^2}{2} = \frac{v_0^2}{2g}.$$

Z rozpatrzonego przykładu widać, że aby jednoznacznie wyznaczyć ruch PM trzeba było podać dwa warunki początkowe, dla $x(t_0)$ oraz dla $dx(t_0)/dt$. Wynika to stąd, że równania Newtona są równaniami różniczkowymi drugiego rzędu i rozwiązanie ogólne zawiera dwie stałe, które wyznacza się (znajduje rozwiązanie szczególne) podając dwa warunki początkowe.

Zapiszmy teraz to samo zagadnienie w notacji wektorowej. Nie zakładamy tu, że wektor przyspieszenia ziemskiego skierowany jest wzdłuż jednej z osi kartezjańskiego układu współrzędnych. Mamy wtedy

$$\vec{F} = -m\vec{g}, \quad \vec{g} = \vec{const}$$

i stąd równanie Newtona

$$\frac{d^2\vec{r}}{dt^2} = -\vec{g}.$$

Warunki początkowe przyjmiemy w postaci

$$\vec{r}(t_0) = \vec{r}_0, \quad \frac{d\vec{r}(t_0)}{dt} = \vec{v}_0$$

i po scałkowaniu równania Newtona dostajemy

$$\frac{d\vec{r}(t)}{dt} - \frac{d\vec{r}(t_0)}{dt} = -\vec{g}(t - t_0)$$

czyli

$$\vec{r}(t) = \vec{r}(t_0) + \frac{d\vec{r}(t_0)}{dt}(t - t_0) - \frac{\vec{g}}{2}(t - t_0)^2 =$$
$$= \vec{r}_0 + \vec{v}_0(t - t_0) - \frac{\vec{g}}{2}(t - t_0).$$

Można to też zapisać w postaci

$$\vec{r}(t) - \vec{r}(t_0) = \vec{v}_0(t - t_0) - \frac{\vec{g}}{2}(t - t_0).$$

Ze wzoru tego widać, że wektor $\vec{r} - \vec{r}_0$ leży w tej samej płaszczyźnie co wektory $\vec{v}$ i $\vec{g}$. Ponieważ wektory te są stałe, więc wniosek stąd taki, że ruch pod wpływem siły przyciągania ziemskiego jest płaski, tzn leży cały czas w płaszczyźnie wyznaczonej przez wektory prędkości początkowej i działającej siły. Jest to szczególny przypadek pewnej ogólnej zasady, którą poznamy później.

Jako drugi przykład całkowania równań Newtona rozpatrzymy oscylator harmoniczny. Jest to najczęściej chyba stosowany model procesu fizycznego. Drgania harmoniczne to ruch jaki wykonuje PM, na który działa siła proporcjonalna do wychylenia z położenia równowagi. Siłę tę możemy zapisać w postaci

$$\vec{F} = -k^2\vec{r},$$

gdzie k jest pewną stałą. Drgania harmoniczne wykonuje np. sprężyna przy zaniedbaniu tarcia. Dla prostoty rozpatrzmy przypadek oscylatora liniowego, tzn jedno-wymiarowego (1D), gdy ruch odbywa się wzdłuż jednej prostej. Wtedy

$$F = -k^2 x$$

i równanie Newtona ma postać

$$m\frac{d^2x}{dt^2} = -k^2 x,$$

co można zapisać jako

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \quad \text{gdzie} \quad \omega^2 = \frac{k^2}{m}. \tag{2.7}$$

Jest to równanie różniczkowe zwyczajne rzędu 2, jednorodne. Aby znaleźć rozwiązanie ogólne takiego równania, dokonujemy podstawienia

$$x = e^{\alpha t}$$

i po wstawieniu do naszego równania, otrzymujemy równanie charakterystyczne (algebraiczne, nie różniczkowe) dla α:

$$\alpha^2 + \omega^2 = 0 \quad \Rightarrow \quad \alpha_{1,2} = \pm i\omega.$$

Tak więc rozwiązanie ogólne, które ma być kombinacją dwu liniowo niezależnych rozwiązań szczególnych (tzn dla α_1 i α_2), ma postać

$$x(t) = A_1 e^{i\omega t} + A_2 e^{-i\omega t}.$$

Korzystając ze wzorów Eulera

$$e^{\pm ia} = \cos a \pm i \sin a,$$

można to zapisać w postaci

$$x(t) = (A_1 + A_2)\cos\omega t + i(A_1 - A_2)\sin\omega t = B_1 \cos\omega t + B_2 \sin\omega t. \tag{2.8}$$

Podstawiając dalej

$$B_1 = h\sin\gamma, \quad B_2 = h\cos\gamma$$

dostajemy

$$x(t) = h\sin(\omega t + \gamma). \tag{2.9}$$

Dowolne dwie stałe, h i γ wyznaczamy z warunków początkowych (położenie i prędkość w chwili początkowej) Jak widać z otrzymanego równania (2.9), ruch pod wpływem siły sprężystej jest okresowy, przy czym okres ten wynosi

$$T = \frac{2\pi}{\omega},$$

bo

$$x(t+T) = h\sin(\omega(t+T)+\gamma) = h\sin(\omega t + 2\pi + \gamma) = h\sin(\omega t + \gamma) = x(t).$$

Dlatego też wielkość ω nazywamy częstością kołową, $1/T \equiv \nu$- częstością, h - amplitudą, argument γ - fazą. Wyznaczmy teraz stałe z warunków początkowych, które mają ogólną postać: $x(0) = x_0$, $dx(0)/dt = v_0$. Różniczkując (2.8) mamy

$$\frac{dx(t)}{dt} = -B_1\omega\sin\omega t + B_2\omega\cos\omega t.$$

Stąd i z (2.8), dostajemy, kładąc $t = 0$

$$B_1 = x_0, \quad B_2 = \frac{v_0}{\omega}$$

i wobec tego rozwiązanie ogólne ma postać

$$x(t) = x_0\cos\omega t + \frac{v_0}{\omega}\sin\omega t. \tag{2.10}$$

Jeżeli w chwili początkowej PM znajdował się w początku układu, czyli w centrum siły sprężystej ($x_0 = 0$) oraz spoczywał ($v_0 = 0$), wówczas z rozwiązania (2.10) wynika, że $x(t) = 0$, a więc PM w dalszym ciągu pozostanie w spoczynku. Możemy więc powiedzieć, że centrum siły sprężystej jest położeniem równowagi oscylatora harmonicznego. Wszystkie te wnioski pozostają słuszne również dla oscylatora 2D i 3D.

1. Zasady zachowania

Klasa równań różniczkowych, które da się rozwiązać ściśle jest niestety niewielka. Dlatego m.in. staramy się zdobyć możliwie dużo informacji o rozpatrywanym układzie fizycznym bez konieczności rozwiązywania równań różniczkowych. Wiele informacji można otrzymać badając symetrie układu. Przejawiają się one w zasadach zachowania. Są to prawa mówiące jakie warunki muszą być spełnione aby pewne wielkości były niezmienne podczas danego procesu. Mówimy wtedy, że wielkości te są zachowane. Z II prawa Newtona wynika bezpośrednio, że

$$\vec{F} = \frac{d\vec{p}}{dt},$$

jeśli więc wypadkowa wszystkich sił działających na PM znika, to

$$\frac{d\vec{p}}{dt} = 0 \Rightarrow \vec{p} = \overrightarrow{const}. \tag{2.11}$$

a więc pęd w czasie procesu jest zachowany, a otrzymany warunek nosi nazwę **zasady zachowania pędu**. Dla otrzymania następnej zasady zachowania wykorzystajmy raz jeszcze równania Newtona (2.2) (narazie są to jedyne znane nam prawa mechaniki) i pomnóżmy je lewostronnie przez $\vec{r}$. Otrzymamy

$$\vec{r} \times \frac{d(m\vec{v})}{dt} = \vec{r} \times \vec{F}. \tag{2.12}$$

Ponieważ jednak $(d\vec{r}/dt) \times (md\vec{r}/dt) = 0$, więc

$$\vec{r} \times \frac{d(m\vec{v})}{dt} = \frac{d\vec{r}}{dt} \times m\frac{d\vec{r}}{dt} + \vec{r} \times \frac{d(m\vec{v})}{dt} = \frac{d(\vec{r} \times m\vec{v})}{dt}.$$

Stąd, uwzględniając (2.12) mamy

$$\frac{d(\vec{r} \times m\vec{v})}{dt} = \vec{r} \times \vec{F}$$

Wprowadzając oznaczenia

$$\vec{J} = \vec{r} \times m\vec{v}, \quad \vec{D} = \vec{r} \times \vec{F}$$

dostajemy

$$\frac{d\vec{J}}{dt} = \vec{D}.$$

Wektor $\vec{J}$ nazywamy **momentem pędu** (krętem), zaś $\vec{D}$ - **momentem siły** (lub momentem obrotowym). Wektor $\vec{J}$ ma kierunek prostopadły do wektorów $\vec{r}$ i $\vec{v}$ (z definicji). Jeżeli więc podczas ruchu kierunek wektora $\vec{J}$ nie ulega zmianie, wówczas ruch jest płaski. Jeżeli moment sił działających na PM znika, tzn $\vec{D} = 0$, wówczas

$$\vec{J} = \overrightarrow{const}. \tag{2.13}$$

i otrzymaliśmy **zasadę zachowania momentu pędu**. Wzór (2.13) nazywamy także **całką pierwszą równań Newtona** lub **całką momentu pędu**. Nazwa całka pierwsza pochodzi stąd, że aby ten wzór otrzymać należy raz scałkować równanie Newtona. Moment sił działających na PM może znikać bądź gdy $\vec{r} = 0$, bądź gdy $\vec{F} = 0$, lub też gdy $\vec{r} \neq 0$, $\vec{F} \neq 0$ lecz $\vec{r} \| \vec{F}$. Taką siłę, która skierowana jest zawsze wzdłuż promienia wodzącego nazywamy **siłą centralną**. Stały punkt do, lub od którego skierowana jest ta siła nazywamy **centrum siły**. Ponieważ dla sił centralnych moment sił $\vec{D} = 0$, więc moment pędu jest zachowany. Stały jest więc kierunek wektora $\vec{J}$, czyli ruch pod działaniem siły centralnej jest płaski. Siłę centralną możemy więc zapisać w postaci

$$\vec{F} = \frac{\vec{r}}{r}F_r, \quad \text{gdzie} \quad F_r = |\vec{F}|, \quad r = |\vec{r}|.$$

Trzecią zasadą zachowania jaką wyprowadzimy będzie zasada zachowania energii. Punktem wyjścia są znów równania Newtona. Tym razem pomnożymy je skalarnie przez prędkość :

$$m\frac{d\dot{\vec{r}}}{dt}\dot{\vec{r}} = \vec{F}\,\dot{\vec{r}}$$

lub

$$\vec{v}m\frac{d\vec{v}}{dt} = \vec{F}\,\dot{\vec{r}}.$$

Lewą stronę można łatwo przekształcić, bo

$$\vec{v}\,\frac{d\vec{v}}{dt} = \frac{1}{2}\,\frac{d\vec{v}^{\,2}}{dt},$$

i wobec tego oznaczając

$$T = \frac{1}{2}m\vec{v}^{\,2},$$

możemy napisać

$$\frac{dT}{dt} = \vec{F}\,\vec{v}. \tag{2.14}$$

Wielkość T nazywamy **energią kinetyczną** PM. Jest ona wyznaczona z dokładnością do stałej addytywnej, niezależnej od czasu, którą tutaj przyjmujemy za równą 0, tak aby w spoczynku $T = 0$, zgodnie z przyjętą umową, że energia kinetyczna jest to energia związana z ruchem ciała. Równanie (2.14) możemy scałkować, otrzymując

$$T - T_0 = \int_{t_0}^{t} \vec{F}\,\vec{v}\,dt = \int_{\vec{r}_0}^{\vec{r}} \vec{F}\,d\vec{r}$$

jako, że $d\vec{r}/dt = \vec{v}$. Wyrażenie $\int_{\vec{r}_0}^{\vec{r}} \vec{F}\,d\vec{r}$ przedstawia pracę wykonaną przez siłę $\vec{F}$ na drodze od $\vec{r}_0$ do $\vec{r}$, gdzie $\vec{r}_0$ to położenie PM w chwili t_0, zaś $\vec{r}$ to położenie w chwili t. Wartość tej pracy zależy na ogół od drogi całkowania. Jeżeli jednak istnieje jednoznaczna funkcja $V(\vec{r}, t)$, taka że

$$\vec{F} = -\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right) \equiv -grad\,V, \tag{2.15}$$

wówczas mówimy, że siła $\vec{F}$ jest **potencjalna**, a samą funkcję $V(\vec{r}, t)$ nazywamy **potencjałem** tej siły. Ponieważ potencjał nie zależy od prędkości, więc też i siła potencjalna może być co najwyżej funkcją położenia i czasu. Jeżeli potencjał jest tylko funkcją położenia, $V = V(\vec{r})$, wówczas nazywamy go **energią potencjalną** PM, zaś o sile (2.15), mówimy, że jest **zachowawcza** lub **konserwatywna** lub **potencjalna**.

Jeżeli więc siła jest potencjalna, to

$$\vec{F}\,d\vec{r} = -grad\,V\,d\vec{r} = -\left(\frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz\right).$$

Jeżeli $V = V(\vec{r})$, czyli jest energią potencjalną, to wyrażenie $\vec{F}d\vec{r}$ jest różniczką zupełną dV i mamy

$$T - T_0 = -\int_{\vec{r}_0}^{\vec{r}} dV(\vec{r}) = -(V(\vec{r}) - V(\vec{r}_0)) = -V + V_0.$$

A więc ostatecznie

$$T + V = T_0 + V_0. \tag{2.16}$$

Wielkość $T + V = E$ nazywamy **energią całkowitą** PM. Ze wzoru (2.16) wynika, że wielkość ta nie ulega zmianie w czasie ruchu o ile siły działające na PM są zachowawcze. Otrzymaliśmy w ten sposób **zasadę zachowania energii**. Z definicji widać wprost, że energia potencjalna wyznaczona jest przez

siłę $\vec{F}$ z dokładnością do stałej addytywnej, tym razem niezależnej od położenia. Umożliwia to upraszczanie szeregu zagadnień przez przyjmowanie odpowiedniej wartości tej stałej. Jest to równoważne po prostu przesuwaniu początku skali wg której mierzymy energię potencjalną. Powierzchnie $V(\vec{r}) = const$ nazywamy **powierzchniami ekwipotencjalnymi**. Ze względu na związek $\vec{F} = -grad\,V(\vec{r})$, pole siły $\vec{F}$ jest wszędzie prostopadłe do powierzchni równego potencjału, $grad\,V(\vec{r})$ wskazuje bowiem kierunek zmian pola $V(\vec{r})$. Prostopadłość $\vec{F}$ oraz powierzchni $V(\vec{r}) = const$ można pokazać z zerowania się iloczynu skalarnego $\vec{F}$ oraz przesunięcia $d\vec{r}$ równoległego do powierzchni ekwipotencjalnej. Biorąc bowiem różniczkę zupełną równania powierzchni stałego potencjału $V(\vec{r}) = const$ mamy

$$dV = 0 = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz = grad\,V\,d\vec{r}.$$

Podobnie jak w wypadku momentu pędu, wyrażenie (2.16) nazywamy całką energii. Jest to całka pierwsza równań Newtona gdy $\vec{F} = -grad\,V(\vec{r})$.

Jako przykład znajdowania potencjału obliczymy potencjał jednorodnego pola grawitacyjnego. Rozpatrzmy PM w pobliżu Ziemi, a więc $\vec{g} = \vec{const}$. i siła ma postać $\vec{F} = m\vec{g}$. Ponieważ

$$\vec{F} = -\frac{dV}{d\vec{r}} = -\nabla V, \quad \text{więc} \quad V(\vec{r}) = V(\vec{r}_0) - \int_{\vec{r}_0}^{\vec{r}} \vec{F}\,d\vec{r}.$$

W przypadku siły grawitacyjnej mamy więc

$$V(\vec{r}) = V(\vec{r}_0) - \int_{\vec{r}_0}^{\vec{r}} m\vec{g}\,d\vec{r} = V(\vec{r}_0) - mg\vec{r} + mg\vec{r}_0.$$

Jeżeli oś O_z będzie zgodna z kierunkiem działania siły grawitacyjnej, wówczas problem staje się jednowymiarowy i dla potencjału mamy

$$V(z) = -mgz + const.$$

Stałą możemy przyjąć za równą zeru i ostatecznie energia potencjalna PM w polu grawitacyjnym jest równa

$$V(z) = -mgz.$$

2. Układ punktów materialnych

Osiągnięte do tej pory wyniki można prosto uogólnić na układ wielu PM. Rozróżniać teraz będziemy siły zewnętrzne $\vec{F}_i^z$ oraz wewnętrzne. Te ostatnie to siły z jakimi dany PM działa na pozostałe PM układu. Równanie ruchu (Newtona) dla i-tego PM ma teraz postać

$$\vec{F}_i^z + \sum_j \vec{F}_{ij} = \dot{\vec{p}}_i, \tag{2.17}$$

gdzie $\vec{F}_{ij}$ to siła oddziaływania i-tego PM na j-ty. Ponieważ PM nie oddziałuje sam na siebie, więc $\vec{F}_{ii} = 0$. Do sił wewnętrznych stosuje się III prawo Newtona, stąd sumując (2.17) po wszystkich punktach układu otrzymujemy

$$\sum_i \vec{F}_i^z + \sum_{ij} \vec{F}_{ij} = \sum_i \dot{\vec{p}}_i = \frac{d}{dt}\sum_i \vec{p}_i,$$

a stąd, ponieważ $\sum_{i,j} \vec{F}_{i,j} = 0$ z trzeciego prawa Newtona, mamy

$$\sum_i \vec{F}_i^z = \frac{d}{dt}\sum_i \vec{p}_i,$$

$\sum_i \vec{F}_i^z$ oznacza sumaryczną siłę zewnętrzną działającą na układ PM. Dla układu PM można wyprowadzić, analogicznie jak dla pojedynczego PM, zasady zachowania. Będą one dotyczyć całkowitej masy układu oraz promienia wodzącego środka masy układu.

3. Więzy

Układ dotąd rozpatrywany był zbiorem swobodnych PM, tzn na jego ruch nie były nałożone żadne ograniczenia i zależał on jedynie od przyłożonej siły $\vec{F}$. Na ogół jednak PM (lub układ PM) musi pozostawać na pewnej powierzchni $f(x, y, z, t) = 0$ (dla i-go PM $f(x_i, y_i, z_i, t) = 0$) lub krzywej $f_1(\vec{r}, t) = 0$, $f_2(\vec{r}, t) = 0$. Ponieważ równania te zawierają jawnie czas, dopuszczamy zmienność tych powierzchni, bądź krzywych w czasie. Przyczyny fizyczne powodujące pozostawanie PM w trakcie ruchu na takiej powierzchni nazywamy **więzami**, zaś odpowiednie równania - równaniami więzów. O funkcjach tych zakładamy, że są ciągłe i dwukrotnie różniczkowalne. Jeżeli równania więzów zależą jawnie od czasu, to więzy takie nazywamy *zależnymi od czasu*, jeśli zaś nie zależą - *stacjonarnymi*. Przykładem więzów stacjonarnych jest ruch PM po nieruchomej powierzchni, np kula tocząca się po stole. Przykład więzów zależnych od czasu to korek unoszący się na powierzchni strumienia.

Jeżeli na PM działa pewna siła $\vec{F}$, wówczas z równania Newtona, znając masę tego PM, możemy wyznaczyć przyspieszenie $\vec{a}$. Z drugiej jednak strony, do równań więzów wchodzi promień wodzący PM - $\vec{r}$. Różniczkując go dwukrotnie względem czasu otrzymujemy wyrażenie na przyspieszenie. Może okazać się, że jest ono różne od przyspieszenia wyznaczonego z równania Newtona, i odwrotnie, $\vec{a}$ z równania Newtona nie spełnia równania więzów. Aby usunąć tę sprzeczność należy siłę w równaniu Newtona uzupełnić o dodatkową wielkość $\vec{F}^R$, tak aby przyspieszenie obliczone z rozszerzonego równania Newtona spełniało równania więzów. $\vec{F}^R$ nazywamy **siłą reakcji więzów**. Nieswobodny PM musi więc spełniać równanie Newtona

$$m\vec{a} = \vec{F} + \vec{F}^R$$

oraz równania więzów

$$f(\vec{r}, t) = 0 \quad \text{lub} \quad f_1(\vec{r}, t) = 0 \quad \text{i} \quad f_2(\vec{r}, t) = 0.$$

Więzy można klasyfikować na rozmaite sposoby. Przyjmiemy tutaj następujący podział - jeżeli ograniczenia nakładane przez istnienie więzów dadzą się przedstawić w postaci równości łączących współrzędne i czas, tzn w postaci

$$f(\vec{r}, t) = 0$$

lub

$$f(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N, t) = 0 \quad \text{dla układu } N \text{ PM}, \tag{2.18}$$

wówczas nazywać je będziemy *holonomicznymi*. Przykładem są więzy w ciele sztywnym, gdzie odległości pomiędzy PM są ustalone. Można to zapisać w postaci

$$(\vec{r}_i - \vec{r}_j)^2 = c_{ij}^2.$$

Natomiast te więzy, których nie da się przedstawić w ten sposób nazywamy *nieholonomicznymi*. Przykładem jest ruch PM wewnątrz sfery, co zapisujemy w postaci

$$\vec{r}_i < a^2, \quad \text{gdzie } a \text{ jest promieniem sfery.}$$

Więzy wprowadzają do zagadnień mechaniki dwie trudności. Pierwsza z nich polega na tym, że nie wszystkie współrzędne $\vec{r}_i$ (dla układu PM) są liniowo niezależne, ponieważ związane są równaniami, bądź nierównościami więzów. Druga trudność polega na tym, że więzy wprowadzają nieznane a priori siły - siły reakcji więzów. Może to być np siła oddziaływania kuli na stół. Siłę taką należy dopiero wyznaczyć. Pierwsza trudność zostaje pokonana przez wprowadzenie tzw **współrzędnych uogólnionych**.

We współrzędnych kartezjańskich, jakimi do tej pory posługiwaliśmy się, układ N PM miał $3N$ niezależnych współrzędnych - jeżeli nie było więzów. Mówimy więc, że miał on 3N stopni swobody. Jeżeli na układ nałożonych jest k równań więzów, postaci (2.18), to spośród $3N$ dotychczasowych zmiennych niezależnych, tylko $3N-k$ pozostanie w dalszym ciągu niezależnymi. Układ będzie więc miał teraz tylko $3N-k$ stopni swobody. Musimy pozbyć się tych k zależnych zmiennych wprowadzając $3N-k$ nowych, niezależnych już, zmiennych $q_1, q_2, \ldots, q_{3N-k}$ i wyrażając stare współrzędne $\vec{r}_1, \ldots, \vec{r}_N$ przez nowe. Otrzymujemy wtedy układ równań

$$\begin{aligned} \vec{r}_1 &= \vec{r}_1(q_1, q_2, \ldots, q_{3N-k}, t), \\ &\ldots \\ \vec{r}_N &= \vec{r}_N(q_1, q_2, \ldots, q_{3N-k}, t). \end{aligned} \tag{2.19}$$

Równania te można traktować jako parametryczne przedstawienie zmiennych $\vec{r}_1, \ldots, \vec{r}_N$, przy czym równania więzów nie wchodzą jawnie do (2.19). O funkcjach $\vec{r}_i(q,t)$ zakładamy także, że są ciągłe i dwukrotnie różniczkowalne . Wprowadźmy oznaczenie

$$n = 3N - k$$

na liczbę stopni swobody układu z więzami, lub, co jest równoważne, liczbę niezależnych współrzędnych uogólnionych. Mają one inny charakter niż współrzędne kartezjańskie. Nie można ich podzielić na trzy grupy, tak aby otrzymać wektory, np $\vec{r} = (x, y, z)$. Nie muszą także mieć wymiaru długości.

Rozpatrzmy ruch PM po powierzchni sfery o promieniu R. Możemy ruch ten opisać przy pomocy współrzędnych kartezjańskich x, y, z, ale, jak nietrudno się domyśleć, wygodniej będzie wprowadzić współrzędne sferyczne

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta.$$

Ponieważ taki PM może poruszać się tylko po powierzchni, a więc gdy $r = R$, co jest równaniem więzów, układ ma tylko dwa stopnie swobody i do jego opisu wystarczą dwie niezależne zmienne. W rozpatrywanym przykładzie będą to dwie zmienne kątowe θ, ϕ, które są nowymi współrzędnymi uogólnionymi, a ich wymiar nie jest długością, lecz miarą kątową.

n-wymiarową przestrzeń rozpiętą przez zmienne q nazywamy **przestrzenią konfiguracyjną**. Każdemu położeniu układu jako całości w chwili t odpowiada punkt w tej przestrzeni. Ruch układu jest tam przedstawiony jako pewna krzywa. Jeżeli więzy nie zależą od czasu, wówczas zawsze można tak wybrać współrzędne uogólnione aby czas nie wchodził jawnie do (2.19). W dalszym ciągu zajmować się będziemy wyłącznie więzami holonomicznymi.

Rozpatrzmy teraz drugą trudność wywołaną przez nałożenie więzów, tzn pojawienie się siły reakcji więzów. Wprowadźmy najpierw pojęcie przesunięcia rzeczywistego i wirtualnego. Niech na nasz układ N PM nałożonych będzie k równań więzów

$$f_\alpha(\vec{r}_i, t) = 0 \quad (\alpha = 1, \ldots, k;\ i = 1, \ldots, N).$$

Obliczając różniczkę tego wyrażenia otrzymujemy

$$\sum_{i=1}^{N} \frac{\partial f_\alpha}{\partial \vec{r}_i}\, d\vec{r}_i + \frac{\partial f_\alpha}{\partial t}\, dt = 0 \tag{2.20}$$

lub

$$\sum_{i=1}^{N} \frac{\partial f_\alpha}{\partial \vec{r}_i} \vec{v}_i + \frac{\partial f_\alpha}{\partial t} = 0.$$

Układ wektorów $\vec{v}_i$ nazywamy *prędkościami dopuszczalnymi* ponieważ przesunięcia, których pochodnymi są prędkości, spełniają równania więzów. Na ogół dla danego czasu t i położenia punktów układu istnieje nieskończenie wiele możliwych prędkości zgodnych z więzami. Podczas rzeczywistego ruchu układu

realizuje się jeden z nich. Z kolei układ nieskończenie małych przesunięć

$$d\vec{r}_i = \vec{v}_i\, dt \quad (i = 1, \ldots, N)$$

gdzie $\vec{v}_i$ to prędkości dopuszczalne, nazywamy przesunięciami *dopuszczalnymi*, bądź *rzeczywistymi*. Przesunięcia te spełniają równania więzów (2.20). Rozpatrzmy teraz dwa układy przesunięć rzeczywistych dla tego samego czasu t i tego samego położenia początkowego układu PM

$$d\vec{r}_i = \vec{v}_i\, dt \quad oraz \quad d\widehat{\vec{r}}_i = \widehat{\vec{v}}_i\, dt.$$

Każde z nich spełnia równania więzów. Wobec tego ich różnica

$$\delta\vec{r}_i = d\widehat{\vec{r}}_i - d\vec{r}_i \quad (i = 1, \ldots, N)$$

spełnia równanie jednorodne

$$\sum_{i=1}^{N} \frac{\partial f_\alpha}{\partial \vec{r}_i}\, \delta\vec{r}_i = 0 \quad (\alpha = 1, \ldots, k) \tag{2.21}$$

Wielkości $\delta\vec{r}_i$ nazywamy **przesunięciami wirtualnymi**. W ogólnym przypadku przesunięcia te nie odpowiadają rzeczywistym zmianom położenia PM, lecz są pomocniczą konstrukcją matematyczną. Aby odczytać sens fizyczny obu typów przesunięć zauważmy, że przesunięcia wirtualne (2.21) różnią się od rzeczywistych (2.20) tym, że nie występuje tam człon

$$(\partial f_\alpha / \partial t)\, dt.$$

Dlatego też często mówi się, że przesunięcia wirtualne są to przesunięcia przy zamrożonych, niezmiennych, więzach. Jeśli bowiem ustalimy w równaniu więzów czas, wówczas różniczkując funkcje f_α człon z dt jest równy 0 i oba wyrażenia są identyczne. W przypadku więzów stacjonarnych nie ma różnicy między obu typami przesunięć.

Mamy wobec tego następujący problem - zadane są siły aktywne $\vec{F}_i = \vec{F}_i(\vec{r}_i, \dot{\vec{r}}_i, t)$ działające na punkty układu i dane są zgodne z więzami położenia początkowe $\vec{r}_i{}^0$ oraz prędkości początkowe $\vec{v}_i{}^0$ punktów układu $(i = 1, \ldots, N)$. Należy wyznaczyć ruch punktów układu i siły reakcji więzów $\vec{F}_i^R$. Przy tak postawionym zadaniu pojawia się trudność - należy wyznaczyć $6N$ wielkości skalarnych - x_i, y_i, z_i oraz F_i^{Rx}, F_i^{Ry}, F_i^{Rz} $(i = 1, \ldots, N)$, mając do dyspozycji mniejszą liczbę równań - $3N$ równań Newtona oraz k równań więzów. k oczywiście musi być mniejsze od $3N$. W przypadku $k = 3N$ układ nie ma żadnego stopnia swobody i nie może się poruszać. Tak więc musimy znaleźć dodatkowe $n = 3N - k$ równań. Otrzymamy je jeżeli ograniczymy się do pewnej klasy więzów, które nazywa się często więzami *idealnymi*. Są to takie więzy, dla których suma prac na przesunięciach wirtualnych znika :

$$\sum_{i=1}^{N} \vec{F}_i^R\, \delta\vec{r}_i = 0. \tag{2.22}$$

Spośród $3N$ wielkości $\delta x_i, \delta y_i, \delta z_i$ tylko $3N - k = n$ jest niezależnych. Możemy więc wykorzystać to równanie aby wyrazić przesunięcia zależne przez n niezależnych. Otrzymamy w ten sposób wyrażenie (2.22) w zmiennych niezależnych. Aby było ono spełnione konieczne jest aby zerowały się współczynniki przy niezależnych zmiennych. To daje nam szukanych n równań. Wprowadzenie przesunięć wirtualnych rozszerzyło klasę rozpatrywanych układów.

4. Ogólne równanie dynamiki.

Postarajmy się teraz znaleźć równanie opisujące ruch układu nieswobodnych PM. Równanie to nie powinno zawierać nieznanych sił reakcji więzów. Mamy równania Newtona

$$\sum_i m\vec{a}_i = \sum_i \left(\vec{F}_i + \vec{F}_i^R\right)$$

oraz (2.22). Wyznaczając $\vec{F}^R$ z równań Newtona i wstawiając do (2.22), dostajemy

$$\sum_{i=1}^{N} \left(\vec{F}_i - m_i\vec{a}_i\right) \delta\vec{r}_i = 0, \tag{2.23}$$

zwane czasami **ogólnym równaniem dynamiki**. $\vec{F}_i$ jest tu wypadkową wszystkich sił aktywnych działających na i-ty PM. Równanie (2.23) jest spełnione dla dowolnego ruchu zgodnego z więzami (o co dbają siły reakcji więzów) i odbywającego się pod działaniem sił $\vec{F}$.

Warto zauważyć, że z ogólnego równania dynamiki (2.23) dla więzów stacjonarnych i układu zachowawczego wynika zasada zachowania energii. Mamy bowiem wtedy

$$\delta\vec{r}_i = d\vec{r}_i = \dot{\vec{r}}_i\, dt,$$

$$\sum \left(\vec{F}_i - \frac{d(m_i\dot{\vec{r}}_i)}{dt}\right) \dot{\vec{r}}_i\, dt = \sum \vec{F}_i\, \dot{\vec{r}}_i dt - \frac{d}{dt}\left(\sum \frac{1}{2} m_i \dot{\vec{r}}_i^{\,2}\right) dt = 0$$

ale $\sum \frac{1}{2} m_i \dot{\vec{r}}_i^{\,2} = T$ oraz dla układu konserwatywnego $\sum \vec{F}_i\, \dot{\vec{r}}_i = -dV/dt$ a więc

$$\frac{d}{dt}\,(V+T)\, dt = 0 \quad \Rightarrow \quad V + T = E = const. \tag{2.24}$$

5. Równania Lagrange'a II rodzaju.

Ogólne równanie dynamiki (2.23) nie ma postaci dogodnej dla rachunków. Wykorzystajmy teraz pojęcie współrzędnych uogólnionych. Pozwoli to nam zredukować liczbę równań ruchu do ilości równej liczbie stopni swobody. Zapiszmy (2.23) we współrzędnych uogólnionych, przejdźmy więc od zmiennych $\vec{r}_i$ $(i = 1, \ldots, N)$ do zmiennych q_j $(j = 1, \ldots, n)$, zgodnie z (2.19) : $\vec{r}_i = \vec{r}_i(q_1, \ldots, q_n, t)$. Dla prędkości, otrzymanych przez różniczkowanie względem czasu (2.19) mamy

$$\vec{v}_i = \sum_{j=1}^{n} \frac{\partial\vec{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial\vec{r}_i}{\partial t} \quad (i = 1, \ldots, N). \tag{2.25}$$

Dla przesunięć wirtualnych (branych przy ustalonym czasie)

$$\delta\vec{r}_i = \sum_{j=1}^{n} \frac{\partial\vec{r}_i}{\partial q_j}\, \delta q_j. \tag{2.26}$$

Elementarna praca przy przesunięciach wirtualnych ma postać

$$\delta A = \sum_{i=1}^{N} \vec{F}_i\, \delta\vec{r}_i,$$

ale ze względu na (2.26)

$$\delta A = \sum_{i=1}^{N} \vec{F}_i \sum_{j=1}^{n} \frac{\partial\vec{r}_i}{\partial q_j}\, \delta q_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{N} \vec{F}_i \frac{\partial\vec{r}_i}{\partial q_j}\right) \delta q_j \equiv \sum_{j=1}^{n} Q_j\, \delta q_j,$$

gdzie Q_j nazywamy **siłami uogólnionymi** i definiujemy poprzez siły aktywne jako

$$Q_j = \sum_{i=1}^{N} \vec{F}_i \frac{\partial \vec{r}_i}{\partial q_j}. \tag{2.27}$$

W położeniu równowagi praca jest równa zero, $\delta A = 0$, a więc również

$$\sum_{j=1}^{n} Q_j \, \delta q_j = 0.$$

Ponieważ wielkości q_j są liniowo niezależne, więc w położeniu równowagi $Q_j = 0$. Znikanie sił uogólnionych jest warunkiem koniecznym i dostatecznym na istnienie położenia równowagi. Podobnie jak współrzędne uogólnione nie muszą mieć wymiaru długości, tak i siły uogólnione nie muszą mieć wymiaru siły, jednak iloczyn $Q_j q_j$ musi mieć wymiar pracy.

Udało się nam wyrazić jeden ze składników (2.23) we współrzędnych uogólnionych. Zróbmy teraz to samo z drugim. Ma on postać

$$\sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \, \delta \vec{r}_i.$$

Różniczkując (2.25) względem $\dot{q}_j$ dostajemy

$$\frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} = \frac{\partial \vec{r}_i}{\partial q_j}. \tag{2.28}$$

Natomiast różniczkując (2.25) względem q_l dostajemy

$$\frac{\partial \dot{\vec{r}}_i}{\partial q_l} = \sum_{j=1}^{n} \frac{\partial^2 \vec{r}_i}{\partial q_j \, \partial q_l} \dot{q}_j + \frac{\partial^2 \vec{r}_i}{\partial t \, \partial q_l} = \frac{d}{dt} \frac{\partial \vec{r}_i}{\partial q_l}$$

i wobec tego

$$\sum_i m_i \ddot{\vec{r}}_i \, \delta \vec{r}_i = \sum_i \sum_j m_i \ddot{\vec{r}}_i \frac{\partial \vec{r}_i}{\partial q_j} \, \delta q_j = \sum_j \left(\sum_i m_i \ddot{\vec{r}}_i \frac{\partial \vec{r}_i}{\partial q_j} \right) \delta q_j \equiv \sum_j Z_j \, \delta q_j$$

Obliczmy teraz Z_j

$$Z_j = \sum_i m_i \ddot{\vec{r}}_i \frac{\partial \vec{r}_i}{\partial q_j} = \frac{d}{dt} \sum_i m_i \dot{\vec{r}}_i \frac{\partial \vec{r}_i}{\partial q_j} - \sum_i m_i \dot{\vec{r}}_i \frac{d}{dt} \frac{\partial \vec{r}_i}{\partial q_j} =$$

$$= \frac{d}{dt} \sum_i m_i \dot{\vec{r}}_i \frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} - \sum_i m_i \dot{\vec{r}}_i \frac{\partial \dot{\vec{r}}_i}{\partial q_j} = \frac{d}{dt} \frac{1}{2} \sum_{i=1}^{N} \frac{\partial m_i \dot{\vec{r}}_i^{\,2}}{\partial \dot{q}_i} - \frac{1}{2} \sum_{i=1}^{N} \frac{\partial m_i \dot{\vec{r}}_i^{\,2}}{\partial q_j}.$$

Ponieważ $T = \sum_i \frac{1}{2} m_i \dot{\vec{r}}_i^{\,2}$ jest energią kinetyczną układu, więc

$$Z_j = \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} \quad (j = 1, \ldots, n)$$

Tak otrzymane wartości możemy teraz podstawić do (2.23) otrzymując ogólne równanie dynamiki we współrzędnych uogólnionych

$$\sum_{j=1}^{n} \left(\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j \right) \delta q_j = 0. \tag{2.29}$$

Wielkości δq_j są niezależne, a więc zerować się muszą wszystkie współczynniki. Otrzymujemy stąd układ równań

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j \qquad (j = 1, 2, \ldots, n). \tag{2.30}$$

Równania te, wynikające z ogólnego równania dynamiki, nazywane są **równaniami Lagrange'a II rodzaju**. Występujące w nich wielkości $\dot{q}_j$ nazywamy **prędkościami uogólnionymi**. Do równań tych nie wchodzą jawnie siły reakcji więzów. Gdybyśmy jednak chcieli siły te wyznaczyć. można to zrobić po scałkowaniu równań Lagrange'a i wyznaczeniu z nich zależności $q_j = q_j(t)$. Zależności te należy następnie podstawić do równań

$$\vec{r}_i = \vec{r}_i(q_1, \ldots, q_n, t)$$

i wyznaczyć stąd $\vec{r}_i = \vec{r}_i(t)$. Następnie dwukrotnie różniczkując względem czasu znaleźć $\vec{a}_i = \ddot{\vec{r}}_i$. Nieznane siły reakcji więzów znajdujemy ze związków

$$\vec{F}_i^R = m_i\vec{a}_i - \vec{F}_i \qquad (i = 1, 2, \ldots, n).$$

Tak więc równania Lagrange'a II rodzaju są szukanym rozwiązaniem postawionego uprzednio problemu. Całkując dwukrotnie te równania otrzymujemy rozwiązanie ogólne, a następnie zadając warunki początkowe ($2n$ warunków - n dla położeń i n dla prędkości) możemy wyznaczyć tor PM w przestrzeni konfiguracyjnej. Jednocześnie możemy też wyznaczyć siły reakcji więzów. Równania Lagrange'a II rodzaju (RL2) są układem n równań różniczkowych zwyczajnych drugiego rzędu dla n nieznanych funkcyj q_j zmiennej niezależnej t. Rząd układu jest więc $2n$. Jeżeli teraz siły uogólnione Q_j nie zależą od prędkości uogólnionych, tzn są postaci $Q_j = Q_j(q_1, \ldots, q_n, t)$ $(j = 1, \ldots, n)$, oraz istnieje taka funkcja $V(q_1, \ldots, q_n, t)$, że

$$Q_j = -\frac{\partial V}{\partial q_j} \quad (j = 1, \ldots, n), \tag{2.31}$$

wówczas siły uogólnione nazywamy potencjalnymi, a funkcję V - potencjałem tych sił. Jeżeli więc siły są potencjalne, możemy RL2 przepisać w postaci

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial (T - V)}{\partial q_j} = 0 \quad (j = 1, 2, \ldots, n).$$

ponieważ jednak potencjał nie zależy od prędkości uogólnionych, można to przepisać jako

$$\frac{d}{dt}\frac{\partial (T - V)}{\partial \dot{q}_j} - \frac{\partial (T - V)}{\partial q_j} = 0 \quad (j = 1, 2, \ldots, n).$$

lub też wprowadzając nową funkcję, zwaną **funkcją Lagrange'a** bądź **Lagrangianem**

$$L = T - V, \tag{2.32}$$

możemy zapisać RL2 w postaci

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0, \tag{2.33}$$

gdzie $L(q_j, \dot{q}_j, t) = T(q_j, \dot{q}_j, t) - V(q_j, t)$. Równania (2.33) to RL2 dla sił potencjalnych.

Pokażemy teraz, że energia kinetyczna układu stacjonarnego jest funkcją jednorodną stopnia drugiego prędkości uogólnionych.

Funkcję $f(x_j)$ nazywamy *funkcją jednorodną stopnia* n zmiennych $x_1, ..., x_M$ jeżeli

$$f(\lambda x_j) \;=\; \lambda^n f(x_j), \quad \sum_{j=1}^{M} \frac{\partial f}{\partial x_j} x_j \;=\; nf \tag{2.34}$$

gdzie λ jest dowolna stałą. Podobnie jest dla funkcji wielu zmiennych. Wyrażenie

$$f(x,y) \;=\; x^2 + 2xy + 3y^2$$

jest funkcją jednorodną stopnia drugiego ponieważ

$$f(\lambda x, \lambda y) \;=\; \lambda^2 x^2 + \lambda^2 2xy + \lambda^2 3xy \;=\; \lambda^2\left(x^2 + 2xy + 3y^2\right) = \lambda^2 f(x,y)$$

Dla energii kinetycznej mamy

$$T = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\vec{r}}_i{}^2 = \frac{1}{2}\sum_{i=1}^{N} m_i \left(\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial \vec{r}_i}{\partial t}\right)^2 =$$

$$= \frac{1}{2}\sum_{i=1}^{N} m_i \left[\left(\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j\right)^2 + 2\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\frac{\partial \vec{r}_i}{\partial t}\dot{q}_j + \left(\frac{\partial \vec{r}_i}{\partial t}\right)^2\right] =$$

$$= \frac{1}{2}\sum_{i=1}^{N} m_i \left[\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\frac{\partial \vec{r}_i}{\partial q_k}\dot{q}_j\dot{q}_k + 2\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\frac{\partial \vec{r}_i}{\partial t}\dot{q}_j + \left(\frac{\partial \vec{r}_i}{\partial t}\right)^2\right] =$$

$$= \frac{1}{2}\sum_{j,k}^{n} a_{jk}\dot{q}_j\dot{q}_k + \sum_{j=1}^{n} a_j\dot{q}_j + a_0$$

gdzie

$$a_{jk} = \sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial q_j}\frac{\partial \vec{r}_i}{\partial q_k}, \quad a_j = 2\sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial q_j}\frac{\partial \vec{r}_i}{\partial t}, \quad a_0 = \frac{1}{2}\sum_{i=1}^{N} m_i \left(\frac{\partial \vec{r}_i}{\partial t}\right)^2.$$

Ogólnie więc energia kinetyczna jest wielomianem stopnia drugiego względem prędkości uogólnionych

$$T = T_2 + T_1 + T_0,$$

$$T_2 = \frac{1}{2}\sum_{j,k} a_{jk}\dot{q}_j\dot{q}_k, \quad T_1 = \sum_{j} a_j\dot{q}_j, \quad T_0 = a_0.$$

Jeżeli nasz układ jest stacjonarny, wówczas więzy i równania łączące $\vec{r}_i$ z q_j nie zależą jawnie od czasu i $\partial r/\partial t = 0$ i wobec tego $a_0 = a_j = 0$. Ostatecznie dla układu stacjonarnego mamy

$$T = T_2 = \frac{1}{2}\sum_{j,k} a_{jk}\dot{q}_j\dot{q}_k. \tag{2.35}$$

Energia kinetyczna jest, jak widać, funkcją jednorodną stopnia 2 prędkości uogólnionych.

Jako pierwszy przykład posługiwania się RL2 rozpatrzmy ruch PM bez więzów. Najprostszy wybór współrzędnych uogólnionych to współrzędne kartezjańskie. Mamy wtedy

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = \frac{\partial T}{\partial z} = 0;$$

$$\frac{\partial T}{\partial \dot{x}} = m\dot{x}, \quad \frac{\partial T}{\partial \dot{y}} = m\dot{y}, \quad \frac{\partial T}{\partial \dot{z}} = m\dot{z}.$$

Na PM działa siła $\vec{F} = (F_x, F_y, F_z)$ i wobec tego RL2 mają postać

$$\frac{d}{dt}(m\dot{x}) = F_x, \quad \frac{d}{dt}(m\dot{y}) = F_y, \quad \frac{d}{dt}(m\dot{z}) = F_z$$

i jak widać, pokrywają się z równaniami Newtona dla PM bez więzów. Równań tych jest trzy, ponieważ na PM nie są nałożone żadne więzy, a więc ma on trzy stopnie swobody i posługujemy się trzema współrzędnymi uogólnionymi, które w tym przypadku są wprost współrzędnymi kartezjańskimi. Jeżeli dana jest siła zewnętrzna $\vec{F}$, to całkując otrzymane równania ruchu możemy (przynajmniej w zasadzie) otrzymać równania toru PM, co jest rozwiązaniem postawionego problemu.

W przykładzie, który przed chwilą rozpatrzyliśmy wybór współrzędnych uogólnionych był właściwie automatyczny. Niestety nie istnieje żadna ogólna metoda. która pozwalałaby na najlepszy wybór współrzędnych uogólnionych. Najlepszy tzn taki, który pozwoli rozwiązać dany problem najszybciej i najprościej. Potrzebne jest tutaj pewne wyczucie i wprawa. Np gdy PM znajduje się pod działaniem siły centralnej, wskazany jest wybór współrzędnych biegunowych, gdyż na ogół siła taka nie zależy od kierunku lecz tylko od odległości od centrum siły, tzn $\vec{F} = \vec{F}(r)$, a więc od jednej tylko współrzędnej biegunowej z pary (r, φ), natomiast we współrzędnych kartezjańskich zależy od obu, związanych ze sobą, zmiennych (x, y).

Jako drugi,nieco mniej banalny przykład, rozpatrzymy ruch PM w polu siły grawitacyjnej. Ruch ten wyznaczaliśmy poprzednio korzystając z równań Newtona. Ruch PM odbywa się w polu siły centralnej, jest więc płaski, co stanowi jedno ograniczenie na ruch PM - jedno równanie więzów. Mamy stąd 2 stopnie swobody i jako współrzędne uogólnione możemy wybrać np współrzędne kartezjańskie (x, y). Potencjał ma postać

$$V = V(y) = mgy$$

natomiast energia kinetyczna

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

i stąd funkcja Lagrange'a

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

$$\frac{\partial L}{\partial y} = -mgy, \quad \frac{\partial L}{\partial \dot{y}} = m\dot{y}, \quad \frac{\partial L}{\partial x} = 0, \frac{\partial L}{\partial \dot{x}} = m\dot{x}$$

Stąd równanie dla współrzędnej x

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \frac{d}{dt}m\dot{x} = 0 \Rightarrow \dot{x} = const = C_1 \Rightarrow x(t) = C_1 t + C_2,$$

i dla współrzędnej y

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = \frac{d}{dt}m\dot{y} + mg = 0$$

i stąd

$$\dot{y} = -gt + C_3 \quad \Rightarrow \quad y(t) = -\frac{1}{2}gt^2 + C_3 t + C_4.$$

Wybierając warunki początkowe odpowiadające np swobodnemu spadkowi z punktu $(0, y_0)$ dostajemy

$$\dot{x}(0) = \dot{y}(0) = 0, \quad x(0) = 0, \, y(0) = y_0.$$

Stąd mamy dla stałych

$$\dot{x}(0) = C_1, \quad x(0) = C_2 \Rightarrow C_1 = C_2 = 0,$$
$$\dot{y}(0) = C_3, \quad y(0) = C_4 \Rightarrow C_3 = 0, \, C_4 = y_0.$$

Ostatecznie więc dostajemy równanie dla PM

$$y(t) = -\frac{1}{2}gt^2 + y_0.$$

Jako następny przykład znajdziemy równanie ruchu (RL2) dla wahadła matematycznego. Jest to PM o masie m zawieszony na nieważkiej i nierozciągliwej nici o zadanej długości l, na który działa siła grawitacji. Zaczniemy od znalezienia liczby stopni swobody takiego układu. W zależności od tej liczby będziemy mieli odpowiednią liczbę współrzędnych uogólnionych. Układ złożony jest z 1 PM, a więc liczba stopni swobody $n = 3 - k$, gdzie k to liczba równań więzów. Ruch wahadła odbywa się w stałej płaszczyźnie, np $z = 0$, co daje jedno równanie więzów. Ponieważ PM zawieszony jest na nici o stałej długości, więc $x^2 + y^2 = l^2 = const$. Mamy dwa równania więzów, czyli nasz układ ma jeden stopień swobody ($n = 3 - 2 = 1$). Wprowadzimy więc jedną zmienną uogólnioną. Najlepiej jest w tym celu wybrać kąt φ wychylenia PM od osi Oy. Mamy wtedy

$$x = l\sin\varphi, \quad y = l\cos\varphi, \quad z = 0.$$

Siła grawitacji działająca na PM skierowana jest np w kierunku ujemnej osi Oy, tzn $\vec{F} = (0, mg, 0)$. Potencjał tej siły

$$V = -\int \vec{F}\, d\vec{r} = -mgy = mgl\cos\varphi.$$

Dalej różniczkując względem czasu wzór określający związek między współrzędnymi kartezjańskimi i współrzędnymi uogólnionymi, dostajemy

$$\dot{x} = l\dot{\varphi}\cos\varphi, \quad \dot{y} = -l\dot{\varphi}\sin\varphi$$

Stąd

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2}m(l^2\dot{\varphi}^2\cos^2\varphi + l^2\dot{\varphi}^2\sin^2\varphi) = \frac{1}{2}ml^2\dot{\varphi}^2.$$

Dla funkcji Lagrange'a mamy

$$L = T - V = \frac{1}{2}ml^2\dot{\varphi}^2 + mgl\cos\varphi.$$

$$\frac{\partial L}{\partial \varphi} = -mgl\sin\varphi, \quad \frac{\partial L}{\partial \dot{\varphi}} = ml^2\dot{\varphi}$$

i jedno RL2 , bo jest jedna współrzędna uogólniona

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}} - \frac{\partial L}{\partial \varphi} = ml^2\ddot{\varphi} + mgl\sin\varphi = 0.$$

Ostatecznie

$$l\ddot{\varphi} + g\sin\varphi = 0.$$

Równanie to daje się prosto rozwiązać w wypadku małych wychyleń, tzn wtedy gdy $\sin\varphi \cong \varphi$. Równanie ruchu ma wtedy postać

$$l\ddot{\varphi} + g\varphi = 0,$$

a więc jest takie jak równanie oscylatora harmonicznego.

6. Prawa zachowania we współrzędnych uogólnionych

Rozpatrując dotąd równania ruchu układu PM nie interesowaliśmy się zupełnie problemem rozwiązania tych równań. Niestety klasa równań jakie dają się scałkować, nie jest duża. Jeżeli mamy bowiem układ PM o n stopniach swobody, wówczas opisuje go układ n równań różniczkowych drugiego rzędu. Musimy więc wykonać ogółem $2n$ całkowań i wyznaczyć z warunków początkowych dla q_j i $\dot{q}_j$ $2n$ pojawiających się stałych. Zadanie takie jest bardzo trudne i trzeba szukać innych sposobów aby uzyskać informację o ruchu układu. W wielu przypadkach można stosunkowo łatwo, bądź nawet odrazu, uzyskać całki pierwsze równań ruchu, tzn wyrażenia postaci $F\left(q_1...q_n, \dot{q}_1...\dot{q}_n, t\right) = \text{const}$, czyli funkcje zawierające tylko pierwszą pochodną położenia względem czasu. Całki te dają często cenne informacje o rozpatrywanym układzie. Często mają one postać **zasad zachowania**, czyli informacji o tym jakie wielkości nie ulegają zmianie w trakcie ruchu układu. Z problemem tym spotkaliśmy się już przy omawianiu równań Newtona używając współrzędnych kartezjańskich. Aby te zasady znaleźć w nowej postaci, tzn korzystając ze współrzędnych uogólnionych, rozszerzymy pewne znane pojęcia. Na początek rozpatrzmy układ zachowawczy, a jako współrzędne uogólnione przyjmijmy współrzędne kartezjańskie. Mamy wtedy

$$\frac{\partial L}{\partial \dot{x}_i} \equiv \frac{\partial T}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i} \sum_k \frac{m_k}{2} \left(\dot{x}_k^2 + \dot{y}_k^2 + \dot{z}_k^2\right) = m_i \dot{x}_i = p_{ix},$$

gdzie p_{ix} jest x-ową składową wektora pędu i-go PM. Tak więc pęd (kartezjański) jest pochodną cząstkową funkcji Lagrange'a względem prędkości (kartezjańskiej) Pojęcie to uogólnimy teraz wprowadzając **pęd uogólniony**, zdefiniowany w analogii do poprzedniej definicji jako

$$p_j = \frac{\partial L}{\partial \dot{q}_j}, \qquad (j = 1, ..., n). \tag{2.36}$$

W ten sposób każdej współrzędnej uogólnionej q_j odpowiada pęd uogólniony p_j, zwany często pędem **kanonicznie sprzężonym** ze współrzędną q_j. Jeżeli q_j nie jest współrzędną kartezjańską i nie ma wymiaru długości, wówczas pęd uogólniony p_j nie ma wymiaru zwykłego pędu.

Zdefiniujmy teraz **współrzędną cykliczną** jako taką współrzędną, która nie wchodzi jawnie do funkcji Lagrange'a. Niech będzie to np współrzędna q_k. Dla takiej współrzędnej

$$\frac{\partial L}{\partial q_k} = 0,$$

a więc RL2 redukują się do

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = 0.$$

To jednak ze względu na (2.36) można zapisać w postaci

$$\frac{dp_k}{dt} = 0 \Rightarrow p_k = const. \tag{2.37}$$

Na tej podstawie możemy sformułować następującą zasadę zachowania:
Jeżeli współrzędna q_k jest cykliczna, to odpowiadający jej pęd uogólniony jest zachowany.

Jak widać, zasada zachowania pędu uogólnionego ma postać całki pierwszej równań ruchu i otrzymaliśmy ją bez konieczności całkowania tych równań, ale jako następstwo faktu, że odpowiednia współrzędna uogólniona jest cykliczna. Od wyprowadzonej poprzednio zasady zachowania pędu obecna różni się tym, że jest ona ogólniejsza. Przy jej wyprowadzeniu nie korzystaliśmy z III zasady Newtona, jest więc słuszna także dla sił działających między poruszającymi się ładunkami elektrycznymi. Po wtóre, z zasady zachowania pędu uogólnionego wynika zarówno zasada zachowania pędu (zwykłego) gdy współrzędna uogólniona odpowiada za przesunięcie układu, jak też i momentu pędu, gdy współrzędna ta związana jest z obrotem układu. Związane jest to z szerszą definicją współrzędnej uogólnionej.

Widać stąd, że współrzędne cykliczne odgrywają dużą rolę przy badaniu własności układu. Jeżeli cykliczna jest współrzędna uogólniona opisująca obrót układu, oznacza to, że własności układu nie zależą od kąta obrotu wokół danej osi. Mówimy wtedy, że układ ma symetrię obrotową, lub że jest **niezmienniczy** względem obrotu wokół tej osi. Sytuacja taka prowadzi do zachowania momentu pędu, ściślej, rzutu momentu pędu na tę oś. Mówimy, że układ jest sferycznie symetryczny jeżeli jest on niezmienniczy względem obrotu o dowolną oś. Zachowany jest wówczas pełny moment pędu. Przy niezmienniczości tylko wobec obrotów wokół jednej osi, zachowany jest tylko odpowiedni rzut momentu pędu na tę oś.

Podobnie jest przy przesunięciach. Jeżeli układ jest niezmienniczy wobec przesunięcia w dowolnym kierunku, zachowany jest całkowity pęd, jeżeli układ jest niezmienniczy tylko wobec przesunięć w jednym kierunku, np wzdłuż jednej osi, - zachowany jest rzut całkowitego pędu na ten kierunek (odpowiednia składowa pędu). Dlatego też jeśli wiadomo, że układ ma określoną symetrię, można podać wówczas odpowiednie całki pierwsze równań ruchu, bez konieczności całkowania tych równań.

Pokażemy teraz, że jeżeli czas jest zmienną cykliczną, wówczas prowadzi to do zasady zachowania energii. Zgodnie z założeniem, funkcja Lagrange'a nie zależy jawnie od czasu: $L = L(q, \dot{q})$. Pochodna zupełna funkcji Lagrange'a ma więc postać

$$\frac{dL}{dt} = \sum_{j=1}^{n} \left(\frac{\partial L}{\partial q_j} \frac{dq_j}{dt} + \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} \right).$$

Z RL2 mamy

$$\frac{\partial L}{\partial q_j} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j}$$

i dlatego

$$\frac{dL}{dt} = \sum_{j=1}^{n} \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} \right) \frac{dq_j}{dt} + \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j} \frac{d\dot{q}_j}{dt} = \sum_{j=1}^{n} \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} \right) \frac{dq_j}{dt} + \sum_{j=1}^{n} \left(\frac{\partial L}{\partial \dot{q}_j} \right) \frac{d}{dt} \left(\frac{dq_j}{dt} \right),$$

a więc

$$\frac{d}{dt} \left(L - \sum_{j=1}^{n} \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) = 0. \tag{2.38}$$

Wynika stąd, że wielkość w nawiasie nie zmienia się w czasie, co można zapisać w postaci

$$L - \sum_{j=1}^{n} \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} = const.$$

Stałą tę oznaczamy przez H :

$$H = \sum_{j=1}^{n} \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L. \tag{2.39}$$

Otrzymaliśmy w ten sposób całkę pierwszą RL2. Pokażemy teraz, że tak zdefiniowana wielkość H jest całkowitą energią układu. Ponieważ energia potencjalna nie zależy od prędkości uogólnionych, więc

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j},$$

Jak wiemy, dla układu o więzach stacjonarnych energia kinetyczna jest kwadratową funkcją prędkości uogólnionych. Wykorzystajmy teraz twierdzenie Eulera o funkcjach jednorodnych, (2.34) co pozwoli nam napisać

$$H = \sum_{j=1}^{n} \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L = \sum_{j=1}^{n} \frac{\partial T}{\partial \dot{q}_j} \dot{q}_j - L = 2T - (T - V) = T + V = E. \tag{2.40}$$

Widać stąd, że wielkość H jest całkowitą energią układu potencjalnego, natomiast (2.38) jest zasadą zachowania energii.

ROZDZIAŁ 3

Zasada Hamiltona

Omówione dotąd równania i zasady, takie jak równania Newtona, Lagrange'a lub d'Alemberta są tzw *zasadami infinitezymalnymi* (zasadami nieskończenie małych) lub różniczkowymi, gdyż przy ich wyprowadzeniu rozważaliśmy położenia układu w nieskończenie małych przedziałach czasu. Zawierają one pochodne względem czasu współrzędnych PM. Oprócz zasad różniczkowych istnieją także w mechanice *zasady całkowe*. O ile zasady różniczkowe rozpatrują ruch układu po zadanej krzywej w kolejnych chwilach czasu, to zasady całkowe dotyczą ruchu układu w całym przedziale czasu, wybierając jedną spośród wielu dopuszczalnych krzywych, po których dany ruch może się odbywać.

Rozpatrzmy teraz ruch układu w przestrzeni konfiguracyjnej, tzn przestrzeni utworzonej przez n współrzędnych uogólnionych $q_1, ..q_n$, przy czym czas t jest parametrem. W przestrzeni tej położenie całego układu w danej chwili t wyznaczone jest przez jeden punkt, a ruch układu wyznacza pewną krzywą - trajektorię układu. Ruch jest więc opisany przez funkcję $q_j = q_j(t), (j = 1, ...n)$ i nazwiemy go *ruchem rzeczywistym*. Zakładamy przy tym, że funkcja $q_j(t)$ jest funkcja ciągłą i dwukrotnie różniczkowalną. Oprócz ruchu rzeczywistego możliwe sa także inne ruchy zgodne z więzami, ale tylko z więzami. Ruchy takie nazywane są *ruchami porównawczymi*. Zakładamy tutaj, że wszystkie ruchy, tak rzeczywisty jak i porównawcze, zaczynają się i kończą w tym samym punkcie. Ruchów porównawczych może być bardzo wiele, ale realizowany jest tylko jeden z nich – ruch rzeczywisty zgodny zarówno z więzami jak i z siłami. Więzy np. mogą ograniczać ruch PM (np kulki) do pewnej powierzchni (np podłogi). Każdy ruch po podłodze będzie więc ruchem dopuszczalnym przez więzy, ale pod działaniem konkretnej siły kulka potoczy się w pewien określony sposób – wybrana będzie konkretna trajektoria, czyli ruch rzeczywisty.

Jeżeli ruch PM zadany jest przez

$$q_j = g_j(t), \tag{3.1}$$

to ruch porównawczy możemy opisać przy pomocy zależności $q_j^*(t)$.

Oba te ruchy różnią się od siebie o pewną wielkość $\delta q(t)$, która może zmieniać się w czasie

$$q_j^*(t) = q_j(t) + \delta q_j(t). \tag{3.2}$$

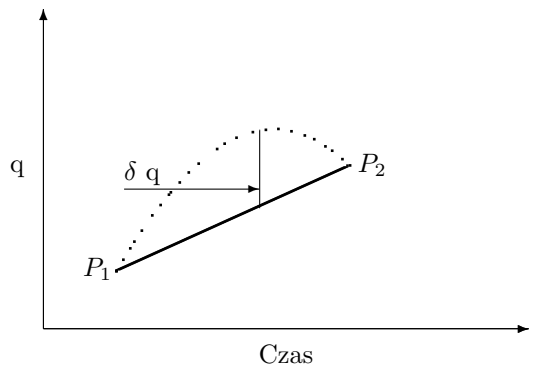
Wariacją współrzędnej uogólnionej nazwiemy teraz wyrażenie $\delta q_j = q_j^*(t) - q_j(t)$, gdzie $q_j^*(t)$ jest ruchem porównawczym. Możemy więc uważać, że wariacja $\delta q_j(t)$ opisuje „odstępstwo" krzywej porównawczej od krzywej dla ruchu rzeczywistego.

Biorąc pochodną względem czasu z $\delta q_j(t)$ dostajemy, że

$$\delta \frac{d}{dt} q_j = \delta \dot{q}_j = \frac{d}{dt} \delta q_j \tag{3.3}$$

oraz podobnie dla całkowania

$$\delta J = \delta \int_{t_0}^{t_1} q(t) dt = \int_{t_0}^{t_1} \delta q(t) dt. \tag{3.4}$$



Trajektoria rzeczywista (linia ciągła) i jedna z trajektorii dopuszczalnych (linia kropkowana). P_1 i P_2 to wspólne punkty – początkowy i końcowy. δq to odstępstwo od ruchu rzeczywistego.

Podobne wzory można też napisać dla współrzędnych kartezjańskich, gdzie teraz wariacja wektora wodzącego będzie miała postać

$$\delta\vec{r}_i = \vec{r}_i - \vec{r}_i{}^*. \tag{3.5}$$

Pamiętając, że $\vec{r}_i = \vec{r}_i(q_1, q_2, ..., q_n)$, możemy wyrazić wirtualną zmianę (wariację) $\vec{r}_i$ poprzez wariację zmiennych q_i

$$\delta\vec{r}_i = \sum_{j=1}^{n} \frac{\partial\vec{r}_i}{\partial q_j}\delta q_j \tag{3.6}$$

a to pokrywa się z podanym poprzednio (2.26) wzorem dla przesunięć wirtualnych. Możemy więc powiedzieć, że wariacje wektorów wodzących są wprost przesunięciami wirtualnymi punktów układu.

W dalszym ciągu interesować się będziemy nie dowolnymi ruchami porównawczymi, lecz tylko takimi dla których położenia w ustalonych chwilach t_0 i t_1 (są to zwykle punkt początkowy i końcowy) pokrywają się. Ruchy te mają więc dwa wspólne punkty w przestrzeni konfiguracyjnej: $q_0 = q(t_0)$ oraz $q_1 = q(t_1)$. W punktach tych wariacje znikają $\delta q_j(t_0) = \delta q_j(t_1) = 0$.

Zasada Hamiltona pozwala znaleźć trajektorię rzeczywistą spośród wszystkich ruchów porównawczych. Punktem wyjścia jest znów ogólne równanie dynamiki (2.23)

$$\sum_{i=1}^{N} \left(\vec{F}_i - m_i\vec{a}_i\right)\delta\vec{r}_i = 0. \tag{3.7}$$

Rozpatrzmy teraz całkę tego wyrażenia

$$\int_{t_0}^{t_1} \sum_{i=1}^{N} \left(\vec{F}_i - m_i\vec{a}_i\right)\delta\vec{r}_i = 0. \tag{3.8}$$

Scałkujmy to przez części wykorzystując fakt, że w punktach początkowym i końcowym wariacje są równe zero, tzn ruchy porównawcze nie różnią się od ruchu rzeczywistego. Dostaniemy, całkując pierwszy wyraz

$$\int_{t_0}^{t_1} m_i \ddot{\vec{r}}_i \delta \vec{r}_i dt = \int_{t_0}^{t_1} m_i \delta \vec{r}_i \frac{d\vec{r}_i}{dt} dt =$$

$$= -\int_{t_0}^{t_1} m_i \dot{\vec{r}}_i \delta \dot{\vec{r}}_i dt = -\int_{t_0}^{t_1} \delta \frac{m_i \dot{\vec{r}}_i^{\,2}}{2} dt. \tag{3.9}$$

Stąd równanie (3.8)

$$-\int_{t_0}^{t_1} \left[\delta \sum_{i=1}^{N} \left(\frac{m_i \dot{\vec{r}}_i^{\,2}}{2} \right) + \vec{F}_i \delta \vec{r}_i \right] dt = 0 \tag{3.10}$$

Ale suma wyrażeń w nawiasie () to całkowita energia kinetyczna T. Możemy więc napisać

$$\int_{t_0}^{t_1} \left[\delta T + \sum_{i=1}^{N} \vec{F}_i \delta \vec{r}_i \right] dt = 0 \tag{3.11}$$

Wzór, który otrzymaliśmy jest to ogólna postać **zasady Hamiltona**, słuszna dla dowolnego typu sił. Załóżmy teraz, że rozpatrywany układ posiada potencjał V, czyli

$$\sum_{i=1}^{N} \vec{F}_i \delta \vec{r}_i = -\delta V \tag{3.12}$$

Zasadę Hamiltona (3.11) możemy teraz zapisać w prostszej postaci

$$\int_{t_0}^{t_1} (\delta T - \delta V)\, dt = 0, \quad \Rightarrow \quad \delta \int_{t_0}^{t_1} (T - V) dt \quad = \delta \int_{t_0}^{t_1} L dt = 0 \tag{3.13}$$

Istnieje wobec tego funkcja Lagrange'a L. Zdefiniujmy wielkość

$$W[q_1 ... q_n] = \int_{t_0}^{t_1} L\left(q_1 ... q_n, \dot{q}_1 ... \dot{q}_n, t\right) dt. \tag{3.14}$$

którą nazywamy **działaniem Hamiltona** w przedziale czasu $[t_0, t_1]$. Działanie to ma wymiar energia $\times$ czas.

Możemy teraz sformułować **zasadę Hamiltona** dla układów potencjalnych:
Dla ruchu rzeczywistego i tylko dla ruchu rzeczywistego, w dowolnym przedziale czasu $[t_0, t_1]$ *mamy*

$$\delta W = \delta \int_{t_0}^{t_1} L\left(q_1 ... q_n, \dot{q}_1 ... \dot{q}_n, t\right) dt = 0. \tag{3.15}$$

Zasada Hamiltona może być punktem wyjścia całej mechaniki klasycznej, można bowiem z niej wyprowadzić RL2.

ROZDZIAŁ 4

Formalizm Hamiltona

1. Równania Hamiltona

Aby wyznaczyć tor PM, na który nałożone są więzy, należy skonstruować funkcję Lagrange'a, a następnie napisać RL2 i rozwiązać je. Spośród rodziny krzywych otrzymanych jako rozwiązanie ogólne, zadając warunki początkowe w postaci $q_j = q_j(t_0)$ i $\dot{q}_j = \dot{q}_j(t_0)$, wyznaczamy tor po jakim w czasie $t > t_0$ poruszać się będzie PM. W ten sposób określamy przyszłość układu. W ten sposób wyrażona jest w mechanice klasycznej zasada przyczynowości. Znajomość równań ruchu oraz warunków początkowych pozwala z całkowitą pewnością odtworzyć przeszłość oraz wyznaczyć przyszłość ruchu układu mechanicznego. W formalizmie Lagrange'a zmiennymi, w liczbie $2n$, są położenia i prędkości uogólnione, równaniami ruchu są równania Lagrange'a II rodzaju, a sam ruch odbywa się w n-wymiarowej przestrzeni konfiguracyjnej zbudowanej na wielkościach $q_1, ..., q_n$.

Możliwy jest i inny opis układu, np przy pomocy współrzędnych uogólnionych q_j i pędów uogólnionych p_j. Używając takich zmiennych przechodzimy do innego formalizmu - formalizmu Hamiltona. Należy teraz znaleźć odpowiednik funkcji Lagrange'a, skonstruować odpowiednie równania ruchu oraz określić przestrzeń w jakiej ruch będzie się odbywał.

Od zmiennych $(q_1, \dot{q}_j, t)$ do zmiennych (q_j, p_j, t) przechodzimy przy pomocy procedury znanej pod nazwą **przekształcenia Legendre'a**. Będzie ona później wielokrotnie stosowana w innych działach fizyki teoretycznej, np w fizyce statystycznej.
Rozpatrzmy dowolną różniczkowalną funkcję $f(x, y)$. Mamy

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \equiv u\, dx + v\, dy.$$

Przejdźmy teraz do nowych zmiennych $(x, y) \to (u, y)$ i wprowadźmy nową funkcję $g(u, y)$

$$g(u, y) = ux - f(x, y).$$

Mamy więc

$$dg = udx + xdu - df \; = \; xdu - vdy$$

a z drugiej strony

$$dg(u, y) = \frac{\partial g}{\partial u} du \; + \; \frac{\partial g}{\partial y} dy,$$

czyli

$$v = -\frac{\partial g}{\partial y}, \quad x = \frac{\partial g}{\partial u}, \quad g(u, y) = ux - f(x, y). \tag{4.1}$$

Wielkości v i x są funkcjami (u, y) zaś związek (4.1) nosi nazwę przekształcenia Legendre'a. Możemy teraz w oparciu o to przekształcenie całkiem formalnie dokonać przejścia od opisu Lagrange'a do opisu Hamiltona. Wystarczy podstawić

$$f \to L, \quad x \to \dot{q}, \quad y \to q. \tag{4.2}$$

Otrzymamy wtedy

$$u = \frac{\partial L}{\partial \dot{q}} = p, \quad v = \frac{\partial L}{\partial q} = \dot{p}. \tag{4.3}$$

Oznaczając nową funkcję g przez H dostajemy

$$H(q,p,t) = \sum_{j=1}^{n} p_j \dot{q}_j - L(q,\dot{q},t). \tag{4.4}$$

Funkcję H, zależącą od wszystkich współrzędnych $q_1, ...q_n$ i pędów uogólnionych $p_1, ..., p_n$, nazywamy **funkcją Hamiltona** lub **Hamiltonianem** układu. Jak łatwo sprawdzić, funkcja H wprowadzona przez przekształcenie Legendre'a i zdefiniowana wzorem (4.4) pokrywa się z wielkością H, którą otrzymaliśmy poprzednio jako całkę pierwszą równań Lagrange'a (wzór(2.40)) Równania ruchu otrzymamy dokonując w (4.1) takiego samego podstawienia (4.2):

$$\dot{p}_j = -\frac{\partial H}{\partial q_j}, \quad \dot{q}_j = \frac{\partial H}{\partial p_j} \qquad (j = 1, ..., n). \tag{4.5}$$

Równania te nazywamy **równaniami Hamiltona** i opisują one ruch układu PM zadanego przez funkcję Hamiltona. Równania te można też otrzymać inaczej. Obliczmy różniczkę funkcji Hamiltona

$$dH = \sum_{j=1}^{n} \left(\frac{\partial H}{\partial q_j} dq_j + \frac{\partial H}{\partial p_j} dp_j \right) + \frac{\partial H}{\partial t} dt. \tag{4.6}$$

Możemy teraz obliczyć różniczkę (4.4) i dostaniemy

$$d\left(\sum_l p_j \dot{q}_j - L(q,\dot{q},t) \right) =$$

$$\sum_j \dot{q}_j dp_j + \sum_j p_j d\dot{q}_j - \sum_j \frac{\partial L}{\partial q_j} dq_j - \sum_j \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j - \frac{\partial L}{\partial t} dt. \tag{4.7}$$

Prędkości uogólnione występujące tutaj należy wyznaczyć przez (q,p,t) przy pomocy relacji odwrotnej do $p = \partial L/\partial \dot{q}$. Relacja taka istnieje, ponieważ jakobian przekształcenia jest różny od zera.
Korzystając ze związków

$$p = \frac{\partial L}{\partial \dot{q}}, \quad \dot{p} = \frac{\partial L}{\partial q}$$

możemy (4.7) sprowadzić do postaci

$$d\left(\sum_l p_j \dot{q}_j - L(q,\dot{q},t) \right) = \sum_j \dot{q}_j dp_j - \sum_j \dot{p}_j dq_j - \frac{\partial L}{\partial t} dt.$$

Porównując to wyrażenie z (4.6) otrzymujemy równania Hamiltona (RH)(4.5) oraz równość

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}, \tag{4.8}$$

wyrażającą fakt, że jeżeli funkcja Lagrange'a nie zależy jawnie od czasu, to również funkcja Hamiltona nie zależy w sposób jawny od czasu.

RH są układem $2n$ równań różniczkowych zwyczajnych rzędu pierwszego. Zamieniliśmy więc układ n równań różniczkowych rzędu drugiego (RL2) na układ $2n$ równań rzędu pierwszego. Z ogólnego punktu widzenia oba formalizmy są prawie równoważne. W mechanice kwantowej i fizyce statystycznej częściej korzysta się z formalizmu Hamiltona.

Z opisem Lagrange'a związana jest n-wymiarowa przestrzeń konfiguracyjna, rozpięta na współrzędnych uogólnionych $q_1...q_n$. Dla formalizmu Hamiltona właściwą przestrzenią jest **przestrzeń fazowa** rozpięta na $2n$ wielkościach - współrzędnych i pędach uogólnionych. Podobnie jak w przestrzeni konfiguracyjnej, czas jest tu parametrem. Ponieważ przestrzeń konfiguracyjna jest n wymiarowa, więc ustalając punkt w tej przestrzeni mamy jeszcze dowolność n zmiennych - prędkości uogólnionych. Natomiast przy ustaleniu punktu w przestrzeni fazowej stan układu jest określony w sposób pełny. Tak więc przez jeden punkt przestrzeni fazowej może, przy zadanych warunkach początkowych, przechodzić tylko jedna krzywa całkowa. Trajektoria w przestrzeni fazowej odpowiada pełnej historii układu.

Metoda Hamiltona jest szczególnie przydatna jeżeli jedna, lub więcej współrzędnych jest cyklicznych. współrzędna taka nie wchodzi do funkcji Lagrange'a, nie wejdzie także do funkcji Hamiltona, gdyż obie te funkcje różnią się tylko o czynnik $\sum p_j \dot{q}_j$, który nie zawiera jawnie q_j. Słuszne jest też twierdzenie odwrotne - jeżeli współrzędna q_k nie wchodzi jawnie do funkcji Hamiltona, to funkcja Lagrange'a również nie zależy jawnie od tej zmiennej. Jest jednak pewna różnica. Jeżeli q_n jest cykliczna, to funkcja Lagrange'a nie zawiera wprawdzie q_n, ale zawiera odpowiednią prędkość uogólnioną. Wiemy, że pęd uogólniony kanonicznie sprzężony ze współrzędną cykliczną jest stały $p_n = const = \alpha$, a więc funkcja Hamiltona ma postać

$$H = H(q_1...q_{n-1}, p_1, ...p_{n-1}, \alpha)$$

i zagadnienie zostało sprowadzone do poziomu $n-1$ stopni swobody. Stałą α wyznaczamy z warunków początkowych. Po wyznaczeniu $n-1$ współrzędnych, współrzędną cykliczną znajdujemy z RH

$$\dot{q}_n = \frac{\partial H}{\partial \alpha}.$$

Do wyznaczenia hamiltonianu można wykorzystać jego związek z funkcją Lagrange'a, jednak często prościej jest skorzystać z faktu, że H jest całkowitą energią układu. Założenia jakie musi spełniać układ aby tak było, są następujące

(1) więzy są stacjonarne (nie zależą od czasu), co sprowadza się do tego, że przekształcenie przeprowadzające współrzędne kartezjańskie w uogólnione nie zależy od czasu,
(2) układ jest zachowawczy, tzn wszystkie siły są potencjalne i potencjał nie zależy od czasu.

Jeżeli spełnione są te warunki, wówczas hamiltonian jest całkowitą energią układu. Ma on zawsze wymiar energii. To, że hamiltonian jest całkowitą energią nie jest równoważne z tym aby był on całką ruchu, a więc aby całkowita energia nie zmieniała się w czasie. Taki warunek ma postać

$$\frac{dH}{dt} = \sum_{j=1}^{n} \left(\frac{\partial H}{\partial q_j} \dot{q}_j + \frac{\partial H}{\partial p_j} \dot{p}_j \right) + \frac{\partial H}{\partial t} =$$

$$= \sum_{j=1}^{n} \left(\frac{\partial H}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial H}{\partial p_j} \frac{\partial H}{\partial q_j} \right) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}. \tag{4.9}$$

Stąd warunek na to aby całkowita energia układu nie zmieniała się w czasie oznacza, że hamiltonian nie może jawnie zależeć od czasu

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}. \tag{4.10}$$

Ze względu na równość (4.8) również funkcja Lagrange'a nie może w takim przypadku zależeć jawnie od czasu.

Przykład

Znajdziemy teraz hamiltonian i RH dla PM poruszającego się po powierzchni kuli o promieniu r i znajdującego się w polu siły ciężkości wytworzonym przez tę kulę. Ze względu na symetrię zagadnienia, najdogodniejszymi współrzędnymi uogólnionymi będą współrzędne kuliste. Mamy jedno równanie więzów

$r = const$, a więc dwa stopnie swobody i dwie niezależne współrzędne uogólnione - dwa kąty.

$$x = r\sin\theta\cos\phi, \qquad \dot{x} = r\dot{\theta}\cos\phi - r\dot{\phi}\sin\theta\sin\phi$$

$$y = r\sin\theta\sin\phi \qquad \dot{y} = r\dot{\theta}\sin\phi = r\dot{\phi}\sin\theta\cos\phi$$

$$z = r\cos\theta, \qquad \dot{z} = -r\dot{\theta}sin\theta.$$

Stąd energia kinetyczna

$$T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) = \frac{1}{2}m\left(r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2\theta\right).$$

Potencjał ma postać

$$V = -\int \vec{F}d\vec{r} = \int m\vec{g}d\vec{r} = mgr + const.$$

Potencjał nie zależy od czasu, jest więc energią potencjalną. Ponieważ każdy potencjał jest wyznaczony z dokładnością do stałej, możemy tę stałą przyjąć jako równą $-mgr$, czyli nasza energia potencjalna będzie równa zeru. Więzy w rozpatrywanym układzie nie zależą od czasu, wszystkie siły są potencjalne, a więc hamiltonian jest energią całkowitą, czyli $H = T = L$, ponieważ $V = 0$. Pędy uogólnione to

$$p_\phi = \frac{\partial L}{\partial\dot{\phi}} = mr^2\phi\sin^2\theta, \quad p_\theta = \frac{\partial L}{\partial\dot{\theta}} = mr^2\dot{\theta}.$$

Ponieważ współrzędna ϕ jest cykliczna, więc p_ϕ jest zachowane. Ze związków tych możemy wyznaczyć prędkości uogólnione jako funkcje pędów

$$\dot{\phi} = \frac{p_\phi}{mr^2\sin^2\theta}, \quad \dot{\theta} = \frac{p_\theta}{mr^2}$$

i podstawić do hamiltonianu otrzymując

$$H = T + V = \frac{1}{2}m\left(r^2\left(\frac{p_\theta}{mr^2}\right)^2 + r^2\sin^2\theta\left(\frac{p_\phi}{mr^2\sin^2\theta}\right)^2\right) =$$

$$= \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2\sin^2\theta}.$$

Mając już funkcję Hamiltona możemy napisać RH

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2}, \quad \dot{p}_\theta = -\frac{\partial H}{\partial\theta} = \frac{p_\phi^2}{mr^2}\frac{\cos\theta}{\sin^3\theta},$$

$$\dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{p_\phi}{mr^2\sin^2\theta}, \quad \dot{p}_\phi = -\frac{\partial H}{\partial\phi} = 0 \Rightarrow p_\phi = const.$$

Jak widać, pierwsze dwie relacje, pochodne kątów względem czasu, nie dały nic nowego. Natomiast drugie pozwalają wyznaczyć zmianę pędów. Pierwsze związki są pożyteczne jeżeli, jak to często ma miejsce w zastosowaniach, znany jest bezpośrednio Hamiltonian, a nie funkcja Lagrange'a.

2. RH z zasady Hamiltona

Wyprowadzimy teraz RH z zasady Hamiltona, którą zapiszemy w postaci

$$\delta \int_{t_0}^{t_1} L dt = \delta \int_{t_0}^{t_1} \left[\sum_{j=1}^{n} p_j \dot{q}_j - H(q,p,t) \right] dt = 0, \tag{4.11}$$

nazywanej czasami *zmodyfikowaną zasadą Hamiltona*. Od poprzedniej różni się tym, że zmiennymi są tu p, q. Obliczmy teraz tę wariację

$$\delta \int_{t_0}^{t_1} \left[\sum_{j=1}^{n} p_j \dot{q}_j - H \right] dt = \int_{t_0}^{t_1} \sum_{j=1}^{n} \left[p_j \delta \dot{q}_j + \dot{q}_j \delta p_j - \frac{\partial H}{\partial q_j} \delta q_j - \frac{\partial H}{\partial p_j} \delta p_j \right] dt =$$

$$= \int_{t_0}^{t_1} \left[\sum_{j=1}^{n} \left(\dot{q}_j - \frac{\partial H}{\partial p_j} \right) \delta p_j + \sum_{j=1}^{n} \left(p_j \frac{d}{dt} - \frac{\partial H}{\partial q_j} \right) \delta q_j \right] dt =$$

$$= \int_{t_0}^{t_1} \left[\sum_{j=1}^{n} \left(\dot{q}_j - \frac{\partial H}{\partial p_j} \right) \delta p_j + \sum_{j=1}^{n} \left(-\dot{p}_j - \frac{\partial H}{\partial q_j} \right) \delta q_j \right] dt + \sum_{j=1}^{n} p_j \delta q_j \Big|_{t_0}^{t_1} =$$

$$= \int_{t_0}^{t_1} \left[\sum_{j=1}^{n} \left(\dot{q}_j - \frac{\partial H}{\partial p_j} \right) \delta p_j + \sum_{j=1}^{n} \left(-\dot{p}_j - \frac{\partial H}{\partial q_j} \right) \delta q_j \right] dt.$$

Ponieważ wariacje δp i δq są niezależne, więc aby całe wyrażenie znikało, jak to wynika ze zmodyfikowanej zasady Hamiltona, muszą zerować się wyrażenia w nawiasach kwadratowych, co daje wprost RH.

3. Nawiasy Poissona

Funkcję $f(q,p,t)$ nazywamy całką równań ruchu jeżeli dla dowolnego ruchu układu wielkość ta pozostaje stała, przy czym wartość tej stałej może być różna dla różnych ruchów. W przypadku RL2, które były równaniami różniczkowymi drugiego rzędu, całka pierwsza tych równań miała postać $f(q,\dot{q},t) = const$. RH są równaniami rzędu pierwszego, stąd całka pierwsza tych równań jest jednocześnie całką ruchu. Równań jest $2n$, a więc maksymalna liczba stałych całkowania też jest równa $2n$. Znając te stałe i oznaczając je np przez c_k, możemy rozwiązać równania

$$f_k(q,p,t) = c_k, \qquad (k = 1, 2, .., 2n) \tag{4.12}$$

otrzymując

$$q_j = \varphi_j(c_1, ..., c_{2n}, t), \quad p_j = \psi_j(c_1, ..., c_{2n}, t), \quad (j = 1, 2, .., n) \tag{4.13}$$

Jeżeli znamy stałe c_k, wówczas z (4.13) wyznaczamy ruch układu i zagadnienie jest rozwiązane. Jeżeli znamy tylko część stałych, to i nasza informacja o ruchu układu nie jest kompletna. Im więcej znamy całek ruchu, tym jest ona pełniejsza. Warto więc szukać metod pozwalających znaleźć nowe całki ruchu. Niech $f(q,p,t)$ będzie taką całką. Wówczas

$$df = \sum_j \left(\frac{\partial f}{\partial q_j} dq_j + \frac{\partial f}{\partial p_j} dp_j \right) + \frac{\partial f}{\partial t} dt = 0.$$

Korzystając z RH można to przepisać inaczej

$$\frac{df}{dt} = \sum_j \left(\frac{\partial f}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial H}{\partial q_j} \right) + \frac{\partial f}{\partial t} = 0. \tag{4.14}$$

Wielkość w nawiasie okazała się być na tyle użyteczną, że warto było wprowadzić dla niej specjalną nazwę - jest to **nawias Poissona** wielkości f i H. Ogólnie nawias Poissona (NP) dwu funkcji f i g, zależących od zmiennych q i p, zdefiniowany jest jako

$$[f,g]_{q,p} = \sum_j \left(\frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j} \right). \tag{4.15}$$

Jak można pokazać bezpośrednim rachunkiem, NP nie zależą od tego od jakich zmiennych kanonicznych zależą funkcje g i f, tzn np od q, p czy też Q, P. Dlatego też będziemy opuszczali wskaźniki przy symbolu NP.

Z definicji wynikają bezpośrednio następujące własności NP

$$[f,g] = -[g,f], \quad c[f,g] = [cf,g], \quad [f+h,g] =, [f,g] + [h,g], \quad \frac{\partial}{\partial t}[f,g] = \left[\frac{\partial f}{\partial t}, g\right] + \left[f, \frac{\partial g}{\partial t}\right]. \tag{4.16}$$

Słuszna jest też *tożsamość Jacobiego*

$$[[f,g],h] + [[g,h],f] + [[h,f],g] = 0. \tag{4.17}$$

Podstawiając w (4.15) $f = q_i, g = q_k$ lub odpowiednio p_i, p_k, gdzie indeksy i, k oznaczają konkretne zmienne, dostajemy

$$[q_i, q_k] = 0, \qquad \text{lub} \quad [p_i, p_k] = 0. \tag{4.18}$$

Podstawiając natomiast $f = q_i, g = p_k$ mamy

$$[q_i, p_k] = \delta_{i,k}, \tag{4.19}$$

gdzie $\delta_{i,k}$ to symbol Kroneckera. Związki (4.19) noszą nazwę *podstawowych nawiasów Poissona*, natomiast zmienne spełniające takie NP nazywamy **zmiennymi kanonicznymi**. Jeżeli więc chcemy sprawdzić czy pewne zmienne Q, P są kanoniczne, to należy w tym celu sprawdzić czy spełniają one podstawowe NP.

Wyznaczmy teraz NP składowych kartezjańskich pędu $\vec{r}$ i momentu pędu $\vec{M} = \vec{r} \times \vec{p}$. Mamy (biorąc $\vec{r} = (x_1, x_2, x_3), x = x_1, y = x_2, z = x_3$)

$$[M_1, p_2] = \sum_{i=1}^{3} \left(-\frac{\partial M_1}{\partial p_i}\frac{\partial p_2}{\partial x_i} + \frac{\partial M_1}{\partial x_i}\frac{\partial p_2}{\partial p_i} \right) = -\frac{\partial M_1}{\partial x_2} = -\frac{\partial}{\partial x_2}(x_2 p_3 - x_3 p_2) = -p_3.$$

Podobnie można pokazać, że

$$[M_i, p_i] = 0, \quad (i = 1, 2, 3).$$

3.1. NP i równania ruchu. Aby powiązać NP z równaniami ruchu, skorzystajmy z definicji NP i zamieńmy tam $f \to q_i$, $h \to H$, a następnie $f \to p_i$, $g \to H$. Korzystając z RH dostajemy odpowiednio

$$[q_i, H] = \frac{\partial H}{\partial p_i} = \dot{q}_i, \qquad [p_i, H] = -\frac{\partial H}{\partial q_i} = \dot{p}_i \tag{4.20}$$

Są to równania ruchu wyrażone poprzez NP i przydają się szczególnie w określeniu czy dana wielkość, tutaj q_i i p_i, jest stałą ruchu. Korzystając z (4.14) mamy dla funkcji f

$$\frac{df}{dt} = [f, H] + \frac{\partial f}{\partial t}. \tag{4.21}$$

Jeżeli więc funkcja f nie zależy jawnie od czasu, wówczas

$$\frac{df}{dt} = [f, H] \tag{4.22}$$

Jeżeli NP wielkości f, nie zależącej jawnie od czasu i hamiltonianu jest równy zeru, to wielkość ta jest stałą ruchu. Prawdziwe jest też twierdzenie odwrotne - jeżeli f jest stałą ruchu i nie zależy jawnie od czasu, to jej NP z hamiltonianem znika. Równość (4.22) jest w takim przypadku kryterium tego czy wielkość f jest stałą ruchu.

Jeżeli znamy już kilka stałych ruchu, to następne można otrzymać z
twierdzenia Jacobiego
Jeżeli f i g są całkami równań ruchu, to NP tych wielkości też jest całką ruchu.
Dla dowodu wystarczy wykazać, że ze związków

$$[f, H] = 0, \quad \text{i} \quad [g, H] = 0 \quad \Rightarrow \quad [[f, g], H] = 0.$$

To jednak wynika bezpośrednio z tożsamości Jacobiego. Jeżeli więc mamy dwie całki równań ruchu, to możemy łatwo skonstruować trzecią. Kłopot polega na tym, że twierdzenie to nie gwarantuje, że otrzymana w ten sposób nowa całka ruchu będzie różna od znanych już, nie będzie ich liniową kombinacją, czy też nie będzie wręcz równa zeru.

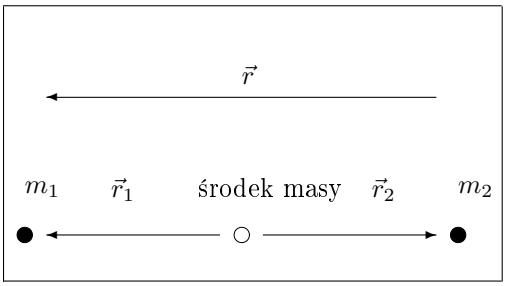
ROZDZIAŁ 5

Zagadnienie dwu ciał

1. Równania toru

Zajmiemy się teraz problemem znalezienia ruchu dwu ciał. Jest to zagadnienie, które, przynajmniej w zasadzie, da się rozwiązać ściśle. Już jednak problem trzech oddziałujących ciał może być rozwiązany tylko numerycznie. Z zagadnieniem dwu ciał mamy do czynienia np rozpatrując w pierwszym przybliżeniu ruch dwu planet. Również słynne doświadczenie rozpraszania cząstek α na jądrach atomowych było opisywane teoretycznie jako zagadnienie dwóch ciał i liczone metodami mechaniki klasycznej.

Rozważmy dwa PM o masach m_1 i m_2. Ze względu na wygodę rachunkową wprowadzimy układ współrzędnych umieszczony w środku mas układu dwu PM. Jeżeli na układ taki nie działają żadne siły oprócz sił oddziaływania pomiędzy PM, wówczas środek mas porusza się ze stałą prędkością, która w naszym układzie wynosi zero. Oznaczmy przez $\vec{r}_1$ i $\vec{r}_2$ położenia PM_1 i PM_2, a przez $\vec{r} = \vec{r}_1 - \vec{r}_2$ wektor wodzący od PM_2 do PM_1.



Rysunek 4. Wektory wodzące w zagadnieniu dwu ciał

Moment pędu i energia kinetyczna mają postać

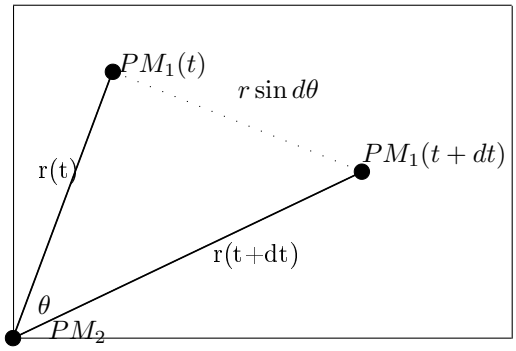
$$\vec{J} = \vec{r}_1 \times m_1\vec{v}_1 + \vec{r}_2 \times m_2\vec{v}_2, \quad T = \frac{1}{2}\left(m_1 v_1^2 + m_2 v_2^2\right). \tag{5.1}$$

$\vec{v}_1, \vec{v}_2$ to odpowiednio prędkości obu PM. Ponieważ

$$\vec{r} = \vec{r}_1 - \vec{r}_2 \quad \text{i} \quad m_1\vec{r}_1 + m_2\vec{r}_2 = 0,$$

bowiem $\vec{r}_1$ i $\vec{r}_2$ wyznaczają położenie punktów materialnych względem środka mas, więc

$$\begin{aligned} \vec{r}_1 &= \frac{m_2}{m_1+m_2}\vec{r}, \qquad & \vec{r}_2 &= \frac{m_1}{m_1+m_2}\vec{r}, \\ \vec{v}_1 &= \frac{m_2}{m_1+m_2}\vec{v}, \qquad & \vec{v}_2 &= \frac{m_1}{m_1+m_2}\vec{v}. \end{aligned} \tag{5.2}$$



Rysunek 5. Współrzędne uogólnione w zagadnieniu dwu ciał.

$\vec{v}$ jest prędkością PM_1 względem PM_2, tak więc PM_2 jest dla nas jak gdyby punktem odniesienia. Podstawiając te wzory do równań dla $\vec{J}, T$ dostajemy

$$\vec{J} = \vec{r} \times \frac{m_1 m_2}{m_1 + m_2} \vec{v} = \vec{r} \times \mu \vec{v},$$

$$T = \frac{1}{2} \frac{m_1 m_2}{m_1 + m_2} v^2 = \frac{1}{2} \mu v^2. \tag{5.3}$$

Wielkość $\mu = m_1 m_2 (m_1 + m_2)^{-1}$ nosi nazwę **masy zredukowanej**.
Tak więc formalnie sprowadziliśmy zagadnienie dwu ciał do problemu ruchu jednego PM o masie μ wokół drugiego PM, uważanego za spoczywający. Oczywiście naprawdę spoczywa tylko środek mas naszego układu. Jeżeli jednak, co często ma miejsce, mamy do czynienia z układem w którym $m_2 \gg m_1$, wówczas środek mas niemal pokrywa się z PM_2, natomiast masa zredukowana jest niemal równa masie lżejszego PM

$$\mu \simeq \frac{m_1}{1 + m_1/m_2} \simeq m_1.$$

Jeżeli na układ działają tylko siły między PM, wówczas moment sił zewnętrznych i wewnętrznych jest równy zero, czyli moment pędu jest stały, a to oznacza że ruch jest płaski. Jako współrzędne uogólnione wybierzmy współrzędne biegunowe (r, θ) położenia PM_1 względem PM_2. Siła działająca między PM będzie skierowana wzdłuż prostej je łączącej. Załóżmy dalej, że siła ta zależy tylko od odległości r, nie zależy natomiast od kierunku. Oczywiście r zależy od czasu jawnie. Zakładamy tym samym, że przestrzeń jest izotropowa i potencjał oddziaływania pomiędzy obu PM zależy tylko od r. Mamy wtedy $V = V(r)$ i nasze zagadnienie ma symetrię kulistą. Funkcja Lagrange'a układu ma postać

$$L = T - V = \frac{1}{2} \mu \left(\dot{r}^2 + r^2 \dot{\theta}^2 \right) - V(r). \tag{5.4}$$

Jak należało oczekiwać, ze względu na symetrię zagadnienia, współrzędna uogólniona θ jest cykliczna. Wobec tego odpowiedni pęd (moment pędu, ponieważ θ jest kątem) jest zachowany

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = \mu r^2 \dot{\theta} = const = |\vec{J}| \equiv \mathcal{J}. \tag{5.5}$$

Jest to całka pierwsza RL2.

Jak widać z rysunku 5, zmiana pola dla małych kątów ma postać $dA = 1/2 r^2 d\theta$. Jeżeli zdefiniujemy prędkość polową PM jako pole zakreślane przez wektor wodzący PM w jednostce czasu, to wyrazi się ona

przez pole dA

$$\frac{dA}{dt} = \frac{1}{2} r^2 \frac{d\theta}{dt}.$$

Łatwo widać, że jedno z RL2 (dla współrzędnej θ) w wypadku sił centralnych, a taka jest siła działająca pomiędzy rozpatrywanymi punktami, wyraża stałość prędkości polowej PM

$$\frac{dA}{dt} = \frac{d}{dt}\left(\frac{1}{2} r^2 \dot{\theta}\right) = 0 \quad \Rightarrow \quad A = const. \tag{5.6}$$

Drugie równanie Lagrange'a (dla współrzędnej r) ma postać

$$\mu \ddot{r} - \mu r \dot{\theta}^2 = F(r), \quad F(r) = -\frac{dV}{dr}. \tag{5.7}$$

Korzystając teraz z całki pierwszego równania Lagrange'a (5.5), możemy w ostatnim równaniu wyeliminować $\dot{\theta}$. Dostaniemy wtedy równanie ruchu w postaci

$$\mu \ddot{r} - \frac{\mathcal{J}^2}{\mu r^3} = F(r), \qquad \dot{\theta} = \frac{\mathcal{J}}{\mu r^2}. \tag{5.8}$$

Możemy teraz zapisać funkcję Lagrange'a, która zależeć będzie od $r, \dot{r}$

$$L = \frac{1}{2}\mu \dot{r}^2 - \frac{\mathcal{J}^2}{2\mu r^2} - V. \tag{5.9}$$

Pierwszy człon to energia kinetyczna, natomiast pozostałe wyrazy (nie zależące od prędkości) to pewien potencjał efektywny

$$V^* = V + \frac{\mathcal{J}^2}{2\mu r^2},$$

który zależy od momentu pędu. Drugi człon to potencjał siły odśrodkowej.

Aby otrzymać równanie toru (5.8) w dogodniejszej formie, zauważmy że drugie równanie da się zapisać jako

$$\mathcal{J} dt = \mu r^2 d\theta$$

i wobec tego dla pochodnych względem t i θ

$$\frac{d}{dt} = \frac{\mathcal{J}}{\mu r^2}\frac{d}{d\theta}$$

i stąd

$$\frac{d^2}{dt^2} = \frac{\mathcal{J}}{\mu r^2}\frac{d}{d\theta}\left(\frac{\mathcal{J}}{\mu r^2}\frac{d}{d\theta}\right).$$

Ostatnią równość możemy wstawić do pierwszego równania (5.8) otrzymując

$$\frac{\mathcal{J}}{r^2}\frac{d}{d\theta}\left(\frac{\mathcal{J}}{\mu r^2}\frac{dr}{d\theta}\right) - \frac{\mathcal{J}^2}{\mu r^3} = F(r). \tag{5.10}$$

Wprowadźmy nową zmienną $x = r^{-1}$ i wykorzystajmy relację

$$\frac{1}{r^2}\frac{dr}{d\theta} = -\frac{d(1/r)}{d\theta} \equiv -\frac{dx}{d\theta}.$$

Otrzymamy w ten sposób wzór na tor w ruchu względnym PM_1 względem PM_2

$$\left(\frac{d^2x}{d\theta^2} + x\right) x^2 = -\frac{\mu}{\mathcal{J}^2} F\left(\frac{1}{x}\right). \tag{5.11}$$

Wzór ten nosi nazwę **wzoru Bineta**. Równanie to przy zadanej sile $F(\frac{1}{x})$ wyznacza orbitę, jeżeli zaś znana jest orbita, pozwala wyznaczyć siłę.

Oprócz całki pierwszej (5.5) jest jeszcze jedna stała ruchu – energia, gdyż układ jest zachowawczy. Energia ma postać

$$E = T + V = \frac{1}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + V(r) = \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\frac{\mathcal{J}}{\mu r^2} + V(r) = const. \tag{5.12}$$

2. Prawa Keplera

Jedną z najczęściej spotykanych sił w przyrodzie (ruch w polu grawitacyjnym lub elektromagnetycznym) jest siła centralna postaci

$$F(r) = -\frac{C}{r^2}, \tag{5.13}$$

gdzie C jest stałą. Jeżeli $C > 0$, to siła jest przyciągająca, bowiem

$$\vec{F} = -C\frac{\vec{r}}{r^3},$$

gdzie $\vec{r} = \vec{r}_1 - \vec{r}_2$ jest wektorem od PM_2 do PM_1. Jeżeli więc rozpatrujemy ruch PM_1 wokół PM_2, ściślej, ruch PM o masie μ wokół środka mas, to dla $C > 0$ siła $\vec{F}$ skierowana jest od PM_1 do PM_2, jest więc siłą przyciągającą. Dla $C < 0$ jest to siła odpychająca.

Potencjał siły $F(r) = -Cr^{-2}$ ma postać

$$V(r) = -\frac{C}{r} = Cx, \qquad x \equiv 1/r, \tag{5.14}$$

a stąd odpowiednie równanie orbity

$$\frac{d^2x}{d\theta^2} + x = \frac{\mu C}{\mathcal{J}^2}. \tag{5.15}$$

Rozwiązaniem tego równania jest

$$x = \frac{1}{r} = \frac{\mu C}{\mathcal{J}^2} + A\cos\theta, \quad \text{a stąd} \qquad r = \frac{p}{1 + \epsilon\cos\theta}. \tag{5.16}$$

Jest to równanie rodziny krzywych stożkowych, a więc otrzymanych przez przecięcie stożka płaszczyznami leżącymi pod różnym kątem. Wielkość ϵ nosi nazwę mimośrodu, zaś p parametru. Jeżeli:

- $\epsilon = 0, \Rightarrow r = \text{const} \Rightarrow$ orbita jest kołem,
- $\epsilon < 1, \Rightarrow r$ jest ograniczone $\Rightarrow$ tor jest elipsą,
- $\epsilon = 1 \Rightarrow 1/r = 0$ w jednym punkcie $\Rightarrow$ tor jest parabolą,
- $\epsilon > 1, \Rightarrow 1/r = 0$ w dwu punktach $\Rightarrow$ tor jest hiperbolą.

Możemy teraz zapisać funkcję Lagrange'a jako

$$L = T - V = \frac{1}{2}\mu\dot{r}^2 + V_{ef}(r), \qquad V_{ef}(r) = \frac{\mathcal{J}^2}{2\mu r^2} - \frac{C}{r}. \tag{5.17}$$

Wprowadziliśmy tu nową wielkość, tzw potencjał efektywny $V_{ef}(r)$, który zależny od r oraz $\mathcal{J}$. Jest to szczególna postać, dla przypadku siły typu siły grawitacyjnej czy coulombowskiej, wprowadzonego przedtem potencjału V^*.

Jeżeli $C < 0$, siła jest odpychająca, funkcja $V_{ef}(r)$ nie osiąga ani minimum ani maksimum w dowolnym przedziale otwartym, czyli orbity są nieograniczone.

Dla sił przyciągania potencjał ten ma jedno minimum, dane przez

$$r_1 = \frac{\mathcal{J}^2}{2\mu r^2}, \quad V_1 = -\frac{\mu C^2}{2\mathcal{J}^2}. \tag{5.18}$$

Całkowita energia ma postać

$$E = \frac{1}{2}\mu\dot{r}^2 + V_{ef} = \frac{1}{2}\mu\dot{r}^2 + \frac{\mathcal{J}^2}{2\mu r^2} - \frac{C}{r}. \tag{5.19}$$

Możliwe są następujące przypadki

(1) $E = V_1$. Istnieje trwała orbita kołowa o promieniu r_1,

(2) $V_1 < E < 0$. Orbity są ograniczone i istnieją dwa punkty zwrotne, wyznaczone przez przyjęcie w wyrażeniu na energię (5.19) $\dot{r} = 0$

$$r_{2,3} = \frac{C}{2|E|} = \pm\sqrt{\frac{C^2}{4E^2} - \frac{\mathcal{J}^2}{2\mu|E|}},$$

(3) $E = 0$. Orbity są ograniczone. Istnieje jeden punkt zwrotny

$$r_4 = \frac{\mathcal{J}^2}{2\mu C},$$

(4) $E > 0$. Orbity nie są ograniczone i istnieje jeden punkt zwrotny

$$r_5 = -\frac{C}{2E} + \sqrt{\frac{C^2}{4E^2} + \frac{\mathcal{J}^2}{2\mu E}}.$$

Jeżeli $C < 0$, co odpowiada sile odpychającej, orbity nie są ograniczone, energia jest zawsze dodatnia i jedyny punkt zwrotny dany jest przez

$$r_6 = -\frac{|C|}{2E} + \sqrt{\frac{C^2}{4E^2} + \frac{\mathcal{J}^2}{2\mu E}}.$$

Jak wiadomo z geometrii analitycznej, półosie elipsy, mała b i duża a, wyrażają się poprzez parametr p i mimośród ϵ następująco

$$a = \frac{p}{1-\epsilon^2}, \qquad b = \frac{p}{\sqrt{1-\epsilon^2}}.$$

Podstawiając tu wartości

$$p = \frac{\mathcal{J}^2}{\mu C}, \quad \epsilon = \sqrt{1 - \frac{2\mathcal{J}^2|E|}{\mu C^2}}$$

dostajemy

$$a = \frac{C}{2|E|}, \qquad b = \frac{\mathcal{J}^2}{\sqrt{2\mu|E|}}. \tag{5.20}$$

Wyznaczyliśmy w ten sposób wielkości charakteryzujące tor po jakim odbywa się ruch względny dwu PM przy zadanych wartościach energii i stałej C. Równanie (5.16) w odniesieniu do ruchu planet wyraża **pierwsze prawo Keplera**, które mówi, że *orbity planet w ruchu wokół Słońca są elipsami, przy czy Słońce znajduje się w jednym z ognisk tej elipsy.*

Pokazaliśmy już poprzednio (5.5), że w przypadku zagadnienia dwu ciał izolowanych od sił zewnętrznych i działających na siebie siłami centralnymi, moment pędu względem środka masy jest stały $\vec{J} = \vec{const}$, bo $\vec{D} = 0$, a więc ruch jest plaski. Pokazaliśmy dalej, że prowadzi to do zachowania wielkości $1/2r^2\dot{\theta}$, zwanej prędkością polową. Fakt ten wyraża **drugie prawo Keplera** (w odniesieniu do ruchu planet), które mówi, że *prędkość polowa w ruchu planet wokół Słońca jest stała*

$$\frac{1}{2}r^2\dot{\theta} = \frac{\mathcal{J}}{2\mu} = const.$$

Ponieważ prędkość polowa jest stała, więc mnożąc ją przez okres ruchu T, otrzymujemy pole elipsy

$$\frac{\mathcal{J}}{2\mu}T = \pi ab.$$

Podstawiając tu teraz wartości dla a i b znalezione w (5.20) dostaniemy

$$T^2 = \frac{2\pi^2\mu a}{|E|} = \frac{4\pi^2\mu}{C}a^3,$$

μ jest masą zredukowaną. Dla sił grawitacyjnych $C = \kappa m_1 m_2$ i mamy

$$T^2 = \frac{4\pi^2}{\kappa(m_1+m_2)}a^3.$$

Jeżeli np $m_2 \gg m_1$, to w mianowniku możemy zaniedbać m_1 wobec m_2 i dostaniemy

$$T^2 \simeq \frac{4\pi^2}{\kappa m_2}a^3. \tag{5.21}$$

Proporcjonalność okresu obiegu planety wokół Słońca do potęgi 3/2 większej półosi elipsy, jest **trzecim prawem Keplera** ruchu planet.

ROZDZIAŁ 6

Szczególna teoria względności

Zajmiemy się teraz badaniem ruchu PM w układzie inercjalnym, przy czym ruch ten odbywać się będzie mógł z prędkością porównywalną z prędkością światła c. Doświadczenia prowadzące do konieczności powstania takiej teorii są znane z kursu podstaw fizyki i nie będziemy ich tu omawiać.

Podstawowym postulatem szczególnej teorii względności jest *równoważność zjawisk fizycznych we wszystkich układach inercjalnych*. Jest to tzw **zasada względności Einsteina**. Można ją sformułować też następująco: *przy pomocy doświadczeń fizycznych nie jesteśmy w stanie rozróżnić „układu spoczywającego" i układu poruszającego się względem niego ruchem jednostajnym*. Zasada Einsteina wymaga więc aby prawa fizyki miały tę samą postać we wszystkich układach inercjalnych. Zasada Einsteina jest uogólnieniem **zasady Galileusza**, która mówi, że prawa mechaniki są identyczne w układach inercjalnych.

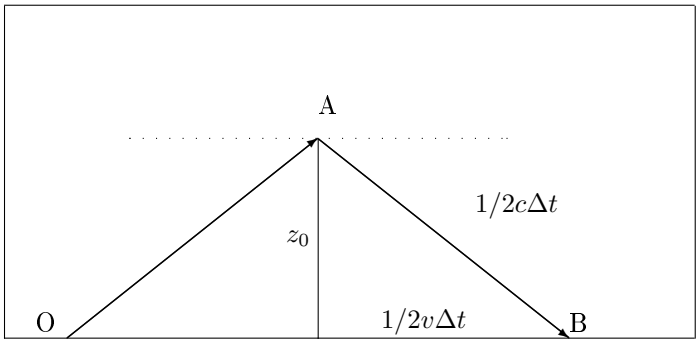
Jeżeli prawa fizyki są takie same we wszystkich układach inercjalnych, to również prędkość światła względem obserwatora jest stała i nie zależy od tego czy obserwator porusza się czy spoczywa. Przyjęcie, za Einsteinem, że prędkość ta jest skończona, co często nazywane jest **drugim postulatem Einsteina**, prowadzi do wielkich zmian obrazu świata.

W mechanice newtonowskiej, gdzie słuszna jest zasada Galileusza, aby przejść od jednego układu współrzędnych do drugiego (zawsze będą to układy inercjalne) posługiwaliśmy się przekształceniem Galileusza

$$\vec{r}' = \vec{r} - \vec{v}t, \quad \Delta t' = \Delta t\,. \tag{6.1}$$

gdzie $\vec{v}$ jest prędkością jednego układu względem drugiego, zaś $\Delta t'$ i Δt są przedziałami czasu mierzonymi odpowiednio w każdym z układów. Jeżeli jednak prędkość światła jest skończona i stała w każdym układzie, to przekształcenie to nie uwzględnia tego faktu. Jeśli bowiem w początku układu U znajduje się źródło światła promieniujące falę kulistą, wówczas prędkość punktu na powierzchni takiej sfery będzie $\dot{\vec{r}} = c\vec{n}$, gdzie $\vec{n}$ jest wersorem skierowanym wzdłuż $\vec{r}$. Różniczkując przekształcenie Galileusza względem czasu otrzymamy $\dot{\vec{r}}' = \dot{\vec{r}} - \vec{v}$, a więc prędkość punktu na powierzchni sfery liczona w układzie U' będzie $\dot{\vec{r}}' = c\vec{n} - \vec{v}$. Będzie ona na ogół różna od c i nie będzie falą kulistą.

Również i prawo transformacji przedziałów czasowych (6.1) nie jest słuszne. Niech bowiem punkt A będzie równo-odległy od punktów B i C w układzie U. Sygnał świetlny wysłany z A dotrze w tym samym czasie do punktów B i C, z których zostanie on w tej samej chwili odesłany z powrotem. W układzie U sygnały te spotkają się w połowie drogi między B i C, a więc w A. Rozpatrzmy teraz tę samą sytuację z punktu widzenia obserwatora w układzie U' (zakładamy odtąd, że układy U i U' poruszają się względem siebie). Przyjmijmy, że w chwili gdy w U do punktów B i C dotrą sygnały, punkty te pokrywają się z punktami B' i C' układu U'. Ponieważ jednak U i U' poruszają się względem siebie, więc sygnały wysłane z B i C nie spotkają się w połowie drogi między B' i C', lecz bliżej jednego z nich (np B'). Prędkość światła jest stała, więc ponieważ w U' sygnały spotkały się nie w połowie drogi, oznacza to, że zostały wysłane nie jednocześnie, lecz sygnał z C' został wysłany wcześniej. Co znaczy, że zjawiska, w naszym przypadku wysłanie sygnałów świetlnych z B i C, które są jednoczesne w jednym układzie, nie są takimi w innym układzie, jeśli układy te poruszają się wzajemnie. Tak więc przekształcenie $\Delta t' = \Delta t$ nie może być spełnione.



Rysunek 7. Droga sygnału w ruchomym układzie współrzędnych.

Trzeba poszukać innego przekształcenia, które zastąpi (6.1). Z drugiej jednak strony cała mechanika newtonowska oparta jest właśnie na tym przekształceniu i wiadomo, że gdy prędkości PM nie są duże, jest ona słuszna. Jeżeli przyjąć, że prędkość światła jest nieskończona to przedstawione powyżej problemy znikają. Uogólnienie wynikające z postulatu Einsteina staje się istotne wtedy gdy prędkości PM są tak duże, że prędkości światła nie można już uważać za nieskończoną. Nowe przekształcenie powinno zachowywać stałą prędkość światła i jednocześnie dla małych prędkości przechodzić w przekształcenie Galileusza. Przekształcenie takie istnieje i nosi nazwę przekształcenia Lorentza (PL).

Równania Newtona są niezmiennicze względem przekształcenia Galileusza, które trzeba będzie zastąpić przez PL. Podobnie może być i dla innych znanych nam równań fizyki. Trzeba więc będzie znaleźć taką postać tych równań, która będzie niezmiennicza względem PL.

Ustalmy najpierw jak będziemy charakteryzowali zdarzenia. Będziemy to robić przez podanie trzech współrzędnych przestrzennych miejsca, względem danego układu współrzędnych, oraz czasu w jakim zdarzenie zaszło. Zakładamy, że w każdym punkcie przestrzeni mamy do dyspozycji zegar. Wszystkie zegary w danym układzie są zsynchronizowane. Można to osiągnąć w ten sposób, że (dla 1D) w punkcie odległym od początku układu o x, początek odczytu czasu ustawiony jest na chwilę $t = x/c$ i uruchamiany przez sygnał świetlny wysłany z początku układu w chwili $t = 0$.

1. Wydłużenie czasu i skrócenie lorentzowskie

Przyjmując zasadę Einsteina można pokazać, że przedział czasu między dwoma zdarzeniami, mierzony w układzie, w którym ciało spoczywa, jest krótszy niż przedział między tymi samymi zdarzeniami mierzony w układzie poruszającym się względem niego. Inaczej mówiąc, przedział czasu mierzony jednym zegarem jest zawsze krótszy niż mierzony dwoma (lub więcej) zegarami.

Rozpatrzmy dwa układy U i U', o względnej prędkości $\vec{v}$. Dla prostoty przyjmijmy, jak to będziemy często dalej robić, że układy poruszają się wzdłuż jednej osi, np Ox. Niech teraz w U' w początku układu O' umieszczony będzie zegar i źródło światła. Na osi prostopadłej do kierunku ruchu, np na osi Oz', w punkcie A i odległości z'_0 od O' znajduje się zwierciadło. Sygnał świetlny wysłany ze źródła dotrze do lustra, odbije się i powróci do punktu O' w czasie $\Delta t' = 2z'_0/c$. Taką sytuację zobaczy obserwator w U'. W U będzie to wyglądało inaczej. Wysłanie i zarejestrowanie sygnału nastąpi nie w jednym, lecz w dwu punktach O i B (p.rysunek 9). Przedział czasu Δt zmierzony będzie w U przy pomocy dwu zegarów. Odległość między punktami O i B wynosi $v\Delta t$. Ponieważ prędkość światła jest stała, więc

trójkąt OAB jest równoramienny i stąd

$$\Delta t = \frac{2}{c}\sqrt{z_0^2 + \left(\frac{v\Delta t}{2}\right)^2}$$

i dalej

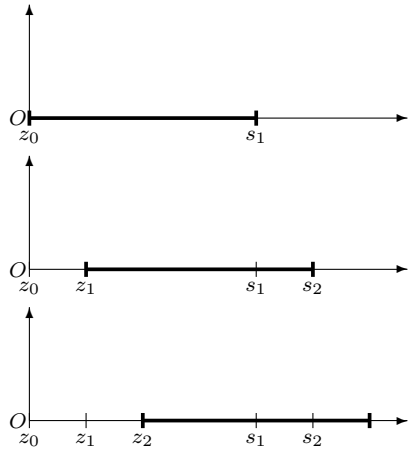
$$\Delta t = \frac{2z_0}{c}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}.$$

Ponieważ długość w kierunku prostopadłym do ruchu nie ulega zmianie, więc $z_0 = z_0'$ i zamiast z_0' możemy podstawić $\frac{1}{2}c\Delta t'$. Ostatecznie dostajemy

$$\Delta t = \frac{\Delta t'}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}. \tag{6.2}$$

W U' oba zdarzenia zaszły w tym samym miejscu, a więc przedział czasu między nimi zmierzony został jednym zegarem. Taki przedział czasu nazywać będziemy przedziałem *czasu własnego* dla tych zdarzeń. Ponieważ pierwiastek w ostatnim wyrażeniu jest nie większy od jedności, więc $\Delta t' \leq \Delta t$, czyli przedział czasu własnego jest najmniejszy.

Zajmiemy się teraz zmianą długości prętów położonych równolegle do kierunku względnego ruch układów. Niech pręt ten spoczywa w U'. Pomiar długości pręta jest tu prosty. Wystarczy podać współrzędne, względem U', początku i końca pręta. Zrobić to można oczywiście w dowolnych dwu, lub jednej chwilach. W układzie U, względem którego pręt się porusza, współrzędne jego końców muszą być oczywiście zmierzone jednocześnie (z punktu widzenia obserwatora w U). Pomiaru dokonajmy w następujący sposób. Lewy koniec pręta umieśćmy w początku U' tzn w O' i tam ustawmy źródło światła z'. Na prawym końcu pręta niech będzie lustro s. Rozpatrzmy dwa zdarzenia - pierwsze to wysłanie z O', w chwili $t' = 0$ sygnału, drugie to zarejestrowanie tego sygnału w O', po jego odbiciu od lustra. Oba te zdarzenia zaszły w jednym miejscu, stąd przedział czasu między nimi to $\Delta t'$ - przedział czasu własnego. Jest on równy $\Delta t' = 2l'/c$, gdzie l' jest długością pręta mierzoną w układzie, w którym on spoczywa. W analogii do czasu, długość tę nazywamy *długością własną*. Wszystkie pomiary z jakimi mamy do czynienia w życiu codziennym dotyczą długości własnej.



Rysunek 8. Pomiar długości w układzie nie związanym z prętem .

Dla obserwatora w U zjawiska te mają inny przebieg. W chwili początkowej $t = t' = 0$ z początku układu U, tzn z $O = O'$, wysłany zostaje sygnał. Lustro znajduje się w położeniu s_1 . Nim jednak sygnał do niego

dotrze, przebędzie ono pewną drogę i znajdzie się w położeniu s_2, zaś źródło w z_1. Sygnał po odbiciu zostanie zarejestrowany w lewym końcu pręta, który będzie w z_2. Względna prędkość układów jest v, możemy więc skorzystać z wyprowadzonego poprzednio wzoru

$$\Delta t = \frac{\Delta t'}{\sqrt{1-\beta^2}}, \quad \text{gdzie} \quad \beta = \frac{v}{c}.$$

Z drugiej strony mamy

$$\Delta t = \frac{z_0 s_2}{c} + \frac{s_2 z_2}{c}$$

Ponieważ lustro przemieszcza się z prędkością v, więc

$$z_0 s_2 = l + v\frac{z_0 s_2}{c} \quad \text{skąd} \quad z_0 s_2 = \frac{l}{1-\beta}$$

$$s_2 z_2 = l - v\frac{z_2 s_2}{c} \quad \text{skąd} \quad s_2 z_2 = \frac{l}{1+\beta}.$$

Ostatecznie

$$\Delta t = \frac{l}{c}\frac{2}{1-\beta^2}, \tag{6.3}$$

gdzie l jest nieznaną narazie (ale stałą) długością pręta mierzoną w U. Porównując teraz oba wzory dla Δt, dostajemy

$$\frac{l'}{\sqrt{1-\beta^2}} = \frac{l}{1-\beta^2} \quad \text{czyli}$$

$$l = l'\sqrt{1-\beta^2}. \tag{6.4}$$

Wzór ten mówi nam jak zmieni się długość pręta jeżeli mierzyć ją będziemy w układzie, w którym pręt spoczywa (l') i w poruszającym się względem pręta. Jak widać, $l \leq l'$, czyli pręt ma największą długość w układzie własnym. Przy przejściu granicznym do mechaniki newtonowskiej, $c \to \infty$, mamy $\Delta t = \Delta t'$ i $l = l'$.

2. Przekształcenie Lorentza

2.1. Przestrzeń pseudoeuklidesowa. Znana przestrzeń euklidesowa ma tę własność, że kwadrat długości wektora x_i

$$x^2 = \sum_i x_i^2$$

jest równy zero tylko wtedy gdy wszystkie składowe wektora są równe zero. Wprowadzając iloczyn skalarny, można to zapisać jako

$$x^2 = (x,x) = \sum_{i,j} x_i x_j (e_i, e_j) = \sum_{i,j} g^{ij} x_i x_j, \tag{6.5}$$

gdzie e_i to składowe wektorów bazy, zaś wielkość

$$g^{ij} = (e_i, e_j)$$

nosi nazwę *tensora metrycznego* gdyż, poprzez zadanie iloczynu skalarnego, określa sposób w jaki w danej przestrzeni mierzona jest odległość.

W trójwmiarowej przestrzeni kartezjańskiej odległość, czyli dystans od początku układu, dana jest poprzez kwadrat długości wektora wodzącego

$$L^2 = \vec{r}^{\,2} = (\vec{r}, \vec{r}) = x^2 + y^2 + z^2 \equiv x_1^2 + x_2^2 + x_3^2 \tag{6.6}$$

tak więc tensor metryczny ma postać

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Jest on diagonalny i na diagonali stoją liczby nieujemne.

Możliwe są jednak i inne przestrzenie, zadane przez inny tensor metryczny. Jeżeli kwadrat długości wektora x da się w tej przestrzeni przedstawić jako suma i różnica kwadratów poszczególnych składowych, to taką przestrzeń nazywamy **pseudoeuklidesową**. Mamy wtedy np

$$(x, x) = x_1^2 + x_2^2 - x_3^2 + \ldots + x_n^2 \,.$$

Jeżeli tensor metryczny przestrzeni pseudoeuklidesowej $\mathcal{R}$ ma na przykład postać

$$g = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

to iloczyn pary wektorów x, y z tej przestrzeni jest równy

$$(x, y) = \sum g^{ik} x_i y_k \; =$$

$$= \; x_1 y_1 (e_1, e_1) + x_2 y_2 (e_2, e_2) + (x_1 y_2 + x_2 y_1)(e_1, e_2) \; = x_1 y_1 - x_2 y_2 \,.$$

Tak też, tzn przy pomocy tensora metrycznego, będziemy dalej zapisywać iloczyn skalarny. Podobnie jak w przestrzeni euklidesowej, dwa wektory nazwiemy ortogonalnymi jeżeli ich iloczyn skalarny znika $(x, y) = 0$.

Przekształcenie liniowe $\mathbf{A}$ przestrzeni pseudoeuklidesowej nazywamy **pseudoortogonalnym** jeżeli zachowuje ono iloczyn skalarny, tzn jeżeli dla $\forall\, x, y \in \mathcal{R}$ mamy

$$(\mathbf{A}x, \mathbf{A}y) = (x, y) \,. \tag{6.7}$$

Przekształcenie takie ma następującą własność

$$\sum_{l,k} g^{lk} a_{il} a_{jk} \; = \; g^{ij}, \tag{6.8}$$

gdzie a_{ij} to elementy macierzowe macierzy przekształcenia A.

Mając wprowadzoną przestrzeń pseudoeuklidesową powróćmy do mechaniki relatywistycznej. Wyprowadzimy teraz przekształcenie Lorentza. Punktem wyjścia jest żądanie aby prędkość światła była stała we wszystkich układach inercjalnych. Równanie czoła fali w układach U i U' winno mieć wobec tego postać

$$c^2 t^2 - x_1^2 - x_2^2 - x_3^2 = 0 \quad \text{oraz} \quad c^2 t'^2 - (x_1')^2 - (x_2')^2 - (x_3')^2 = 0 \,.$$

Przyrównując oba wyrażenia dostajemy

$$c^2 t^2 - \sum_{i=1}^{n} x_i^2 = c^2 t'^2 - \sum_{i=1}^{n} x_i'^2 \,, \quad (x_i = x, y, z) \,. \tag{6.9}$$

Wprowadźmy czwartą składową, rzeczywistą, proporcjonalną do czasu, także o wymiarze długości, i oznaczmy ją przez $x_0 = ct$. Wskaźniki greckie $(\alpha, \beta, \ldots)$ przebiegać będą cztery wartości $0, 1, 2, 3$. Możemy teraz (6.9) zapisać jako

$$x_o^2 - \sum_{i=1}^{3} x_i^2 = x_o'^2 - \sum_{i=1}^{3} x_i'^2 \,.$$

Ale wyrażenie $x_o^2 - \sum_{i=1}^{3} x_i^2$ jest kwadratem długości czterowektora. Możemy je zapisać korzystając z tensora metrycznego jako $\sum g^{\mu\nu} x_\mu x_\nu$, gdzie $g^{\mu\nu}$ ma różne od zera wyrazy tylko na przekątnej i tam

$g^{00} = -g^{11} = -g^{22} = -g^{33} = 1$. Wobec tego warunek (6.9) oznacza, że szukane przekształcenie Lorentza nie powinno zmieniać iloczynu skalarnego w przestrzeni o tensorze metrycznym

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{6.10}$$

Czterowymiarową przestrzeń tak zbudowaną, tzn zawierającą składową proporcjonalną do czasu oraz trzy składowe przestrzenne, nazywamy **czasoprzestrzenią** lub **przestrzenią Minkowskiego**. Jest to przestrzeń pseudoeuklidesowa, a szukane przekształcenie Lorentza jest przekształceniem pseudoortogonalnym. Elementy $a_{\mu\nu}$ tego przekształcenia powinny więc spełniać warunek

$$g^{\mu\nu} = \sum_{\lambda,\rho=0}^{3} a_{\mu\lambda} a_{\nu\rho} g^{\lambda\rho} \,. \tag{6.11}$$

Ponieważ czasoprzestrzeń jest jednorodna, więc związek łączący x_μ i x'_μ powinien być liniowy. Tylko wtedy bowiem zmiana początku układu U prowadzi do takiego samego dla wszystkich punktów przesunięcia początku układu U', co jest konieczne aby układy te były równoważne. Niech względna prędkość obu układów będzie skierowana wzdłuż osi x_1. Przyjmiemy też, że osie układów są wzajemnie równoległe i że początki układów pokrywają się w chwili $t = t' = 0$. Tak więc z $y = 0$ i $z = 0$ powinno wynikać $y' = 0$ oraz $z' = 0$. Ponieważ obroty wokół osi przestrzennych nie powinny zmieniać fizyki, nie będziemy ich tu rozpatrywać i przyjmiemy, że płaszczyzna $y = 0$ przechodzi w płaszczyznę $y' = 0$ a $z = 0$ w $z' = 0$. Oznacza to, że $y' = ay$ oraz $z' = az$. Aby układy były zupełnie równoważne a powinno być równe 1. Macierz PL powinna więc wiązać ze sobą tylko dwie składowe czteroprzestrzeni - czas oraz składową wzdłuż osi której odbywa się ruch (w naszym przypadku x_1):

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} & 0 & 0 \\ a_{10} & a_{11} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.12}$$

Jednocześnie współczynniki macierzy $\mathbf{A}$ powinny spełniać warunki pseudoortogonalności (6.8):

$$g^{00} = \sum_{\rho,\nu=0}^{3} g^{\rho\nu} a_{0\rho} a_{0\nu} = g^{00} a_{00}^2 + g^{01} a_{00} a_{01} + g^{10} a_{01} a_{00} + g^{11} a_{01}^2 + \ldots + g^{33} a_{03}^2 \,,$$

stąd uwzględniając wartości elementów a_{ij} danych przez (6.12), dostajemy

$$g^{00} = a_{00}^2 - a_{01}^2 = 1 \,. \tag{6.13}$$

Podobnie dla g^{11}

$$a_{10}^2 - a_{11}^2 = -1 \,, \tag{6.14}$$

oraz dla g^{10}

$$a_{00} a_{10} - a_{01} a_{11} = 0 \,. \tag{6.15}$$

Ponieważ warunków tych jest za mało aby jednoznacznie określić współczynniki $a_{\mu\nu}$, skorzystamy z równania ruchu środka układu U' (tzn punktu $x'_1 = 0$) względem układu U. Porusza się on ze stałą prędkością v wzdłuż osi Ox_1, a więc po upływie czasu t jego współrzędna w układzie U ma wartość

$$x_1 = vt \equiv \beta x_0, \quad (\beta \equiv \frac{v}{c}) \,.$$

Dlatego też, korzystając z ogólnego wzoru na transformację wektora $\vec{x}$, ($x' = Ax$) początek układu U' może być opisany przez

$$x'_1 = a_{10}x_0 + a_{11}x_1 = (a_{10} + \beta a_{11})x_0 = 0\,,$$

a stąd

$$a_{10} = -\beta a_{11}\,. \tag{6.16}$$

Drugi warunek pseudoortogonalności (6.14) możemy teraz przepisać jako

$$\beta^2 a_{11}^2 - a_{11}^2 = -1\,,$$

a stąd

$$a_{11} = \frac{1}{\sqrt{1-\beta^2}}$$

oraz

$$a_{10} = \frac{-\beta}{\sqrt{1-\beta^2}}\,.$$

Wstawiając otrzymane związki do trzeciego warunku dostajemy

$$a_{00} = \frac{1}{\sqrt{1-\beta^2}}, \quad a_{01} = \frac{-\beta}{\sqrt{1-\beta^2}}\,.$$

Szukane PL ma więc postać

$$\begin{aligned} x'_0 &= \gamma(x_0 + \beta x_1), \\ x'_1 &= \gamma(-\beta x_0 + x_1), \\ x'_2 &= x_2, \\ x'_3 &= x_3, \end{aligned} \tag{6.17}$$

a jego macierz

$$\mathbf{A} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{6.18}$$

gdzie $\gamma = (1-\beta^2)^{-1/2}$. Aby otrzymać macierz przekształcenia odwrotnego należy zamienić v na $-v$, czyli β na $-\beta$.

3. Wnioski z przekształcenia Lorentza (PL)

Jednoczesność. PL jest to sposób na przetłumaczenie współrzędnych zdarzenia z jednego układu współrzędnych do innego. Z postaci (6.17) wynika od razu, że jeżeli układy te poruszają się względem siebie to zdarzenia, które są jednoczesne w jednym układzie nie mogą nimi być w drugim. Rozpatrzmy bowiem dwa zdarzenia jednoczesne w U o współrzędnych $(x_1, y_1, z_1, t_1 = t)$ oraz $(x_2, y_2, z_2, t_2 = t)$. Korzystając z PL (6.17) mamy dla odpowiednich czasów w U':

$$t'_1 = \gamma(t_1 - \frac{v}{c^2}x_1) = \gamma(t - \frac{v}{c^2}x_1)\,,$$

$$t'_2 = \gamma(t_2 - \frac{v}{c^2}x_2) = \gamma(t - \frac{v}{c^2}x_2)\,.$$

Różnica tych czasów

$$\Delta t' = t'_2 - t'_1 = \frac{v}{c^2}(x_1 - x_2) \neq 0 \quad \text{gdy} \quad x_1 \neq x_2\,. \tag{6.19}$$

Jeżeli więc dwa zdarzenia zaszły w różnych punktach, to nie są one jednoczesne w U', mimo że były takie w U. To samo dotyczy synchronizacji zegarów. Zegary zsynchronizowane w U będą niezsynchronizowane w U' i różnica wskazań będzie rosła wraz z odległością między nimi. Przyjmując, że prędkość światła jest nieskończona, otrzymujemy zachowanie jednoczesności.

Zmiana długości poruszającego się pręta. Podobnie korzystając z PL (6.17) można pokazać, że długość własna jest największa. Aby zmierzyć długość pręta w układzie poruszającym się względem niego, należy pomiaru końców pręta dokonać w tej samej chwili. Mamy więc z (6.17)

$$x_2' = \gamma(x_2 - vt), \quad x_1' = \gamma(x_1 - vt), \quad \text{bo} \quad t_1 = t_2 = t\,.$$

Stąd

$$L_0 = x_2' - x_1' = \gamma(x_2 - x_1) = \gamma L,$$

czyli

$$L = L_0\sqrt{1-\beta^2}\,. \tag{6.20}$$

Oczywiście taka zmiana długości pręta nie oznacza, że ulega on rzeczywistemu skróceniu, czyli że przy badaniu go w jednym z układów pojawią się w nim naprężenia. Byłoby to sprzeczne z podstawową zasadą szczególnej teorii względności, że prawa fizyki są takie same we wszystkich układach inercjalnych. Różnica sprowadza się więc do sposobu pomiaru, czy jest on dokonywany przy pomocy jednego czy też więcej zegarów. Jeżeli pomiar dokonywany musi być przy pomocy więcej niż jednego zegara, wówczas nie jest zachowana jednoczesność i mierzona jest inna wartość niż w układzie, gdzie pomiaru dokonuje się przy użyciu jednego zegara. Można też pokazać, opierając się o PL, że jeżeli w U' w punkcie (x_1', y_1', z_1') zaszły dwa zdarzenia, jedno w chwili t_1', drugie w t_2', to w U okaże się, że

$$x_1 = \gamma(x_1' + vt_1'), \quad x_2 = \gamma(x_2' + vt_2'),$$

i jeżeli $t_1' \neq t_2'$, to podobnie jak nie była zachowana jednoczesność, teraz nie jest zachowana jedność miejsca. Jedynie zjawiska, które zachodzą w tym samym czasie i w tej samej chwili pozostają takimi po przejściu do innego układu współrzędnych.

Przyczynowość. Jeżeli w układzie U zaszły dwa zdarzenia, przy czym zdarzenie 1 było przyczyną zdarzenia 2, to oczywiście najpierw powinno zajść 1 a potem dopiero 2. Ten stan rzeczy, tzn związek przyczynowy między dwoma zjawiskami, powinnien zostać zachowany w dowolnym innym układzie, wszystkie one są bowiem równoprawne. Weźmy teraz jedno z równań PL

$$t' = \gamma(t - \frac{v}{c^2}x)$$

i zróżniczkujmy je. Otrzymamy

$$dt' = \gamma(dt - \frac{v}{c^2}dx)\,.$$

Ponieważ $dx/dt = v_x$, więc

$$\frac{dt'}{dt} = \gamma(1 - \frac{v}{c^2}v_x)\,. \tag{6.21}$$

Aby utrzymana była sekwencja zdarzeń w obu układach, stosunek przyrostów czasu musi zachować znak, czyli $dt'/dt > 0$, ponieważ dt to czas między zdarzeniami 1 i 2 mierzony w układzie U, natomiast dt' to czas między tymi samymi zdarzeniami mierzony w układzie U'. Zachowanie znaku oznacza, że w obu układach np. zdarzenie 1 poprzedzało zdarzenie 2. Prędkość względna v obu układów może być dodatnia lub ujemna, lecz zawsze mniejsza od c, tak więc aby zachować znak stosunku przyrostów czasu, v_x musi być mniejsze od c. Oznacza to, że jeżeli dwa zdarzenia mają być powiązane ze sobą związkiem przyczynowym, to prędkość z jaką przesyłany jest sygnał zawierający informację powodującą zależność przyczyna - skutek, musi być mniejsza od prędkości światła.

4. Przedziały między zdarzeniami i ich klasyfikacja

Rozpatrzmy dwa zdarzenia I i II. Dla prostoty rachunków przyjmijmy, że zaszły one na osi Ox układu U, natomiast U' porusza się względem U wzdłuż osi Ox. Współrzędne tych zdarzeń będą miały postać (jeżeli będziemy używać zmiennej ct zamiast x_0)

$$\begin{aligned} I \quad & (ct_1, x_1, 0, 0), \quad (ct_1', x_1', 0, 0), \\ II \quad & (ct_2, x_2, 0, 0), \quad (ct_2', x_2', 0, 0)\,. \end{aligned}$$

Stosując PL (6.17) mamy

$$x_2' - x_1' = \gamma\,[\,(x_2 - x_1) - v(t_2 - t_1)\,]$$

$$t_2' - t_1' = \gamma\left[(t_2 - t_1) - \frac{\beta}{c}(x_2 - x_1)\right]. \tag{6.22}$$

Widać więc, że warunkiem na to aby zachowana była jednoczesność i jedność miejsca, jest aby $x_2 = x_1$ oraz $t_2 = t_1$. Pojęcie przedziału pomiędzy dwoma chwilami lub dwoma punktami w przestrzeni jest względne i wartość jego zmienia się przy przejściu od jednego układu do drugiego. Można postawić pytanie czy wartość ta może się zmieniać dowolnie. Zauważmy najpierw, że istnieje, z założenia, wielkość która jest niezmiennikiem PL. Jest to odległość w przestrzeni Minkowskiego jaką w ustalonym czasie przebędzie sygnał świetlny. Oznaczmy ją przez l_{12}:

$$l_{12}^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2 = 0\,. \tag{6.23}$$

Wielkość związana z odległością w czteroprzestrzeni między dwoma zdarzeniami to

$$s_{12} = \sqrt{c^2\,(t_2 - t_1)^2 - (\vec{r}_2 - \vec{r}_1)^2} \tag{6.24}$$

jest również niezmiennikiem PL i nazywamy ją **przedziałem** między zdarzeniami I i II. W postaci różniczkowej

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 = dx_0^2 - \sum dx_i^2 = \sum g^{\mu\nu}dx_\mu dx_\nu = (dx, dx)\,.$$

Wzór ten można zinterpretować następująco: jeżeli w trójwmiarowej przestrzeni porusza się PM, wówczas w cztero wymiarowej czasoprzestrzeni odpowiadający PM będzie przebiegał pewien tor, który za Minkowskim nazwiemy **linią świata** tego PM. Poszczególne punkty na linii świata to położenia PM w różnych chwilach czasu. Oznaczając przez dx_μ wektor przesunięcia wzdłuż tej linii, możemy powiedzieć, że

$$ds^2 = \sum g^{\mu\nu}dx_\mu dx_\nu = (dx, dx) \tag{6.25}$$

jest kwadratem długości łuku linii świata. Wiemy już, że ds^2 jest niezmiennikiem PL, podczas gdy dt i dx_i nimi nie są. Postarajmy się teraz znaleźć niezmiennik PL proporcjonalny do czasu. W tym celu podzielmy ds^2 przez inny niezmiennik c^2. Oznaczając otrzymaną wielkość przez $d\tau^2$ mamy

$$d\tau^2 = c^{-2}\sum g^{\mu\nu}dx_\mu dx_\nu = c^{-2}(dx, dx)\,. \tag{6.26}$$

Aby znaleźć sens fizyczny wielkości $d\tau$, obliczmy $d\tau^2$ w układzie U', względem którego cząstka spoczywa. Będzie to więc układ poruszający się wraz z cząstką wzdłuż jej linii świata. Wobec tego w U' współrzędne przestrzenne cząstki nie ulegają zmianie, tzn $dx_\mu' = 0$ i cały wektor $dx_\mu' = (c\,dt', 0, 0, 0)$. Stąd

$$d\tau = c^{-2}\sum g^{\mu\nu}dx_\mu' dx_\nu' = (dt')^2\,.$$

Oznacza to, że $d\tau = dt'$, czyli $d\tau$ jest przedziałem czasu mierzonym w układzie w którym cząstka spoczywa, czyli jest to czas własny. Rozpisując definicję $d\tau$ można znaleźć związek między czasem własnym i czasem w innym układzie :

$$d\tau = dt\sqrt{1 - \beta^2}\,.$$

Powróćmy do naszych dwu zdarzeń I i II. Niech w U współrzędne przestrzenne będą odległe o

$$L_{12}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

zaś czasowe o $t_{12} = t_2 - t_1$, Możemy więc napisać

$$s_{12}^2 = c^2 t_{12}^2 - L_{12}^2 \,,$$

przy czym przedział czasoprzestrzenny jest niezmiennikiem PL. Niech

$$c^2 t_{12}^2 > L_{12}^2 \,.$$

Oznacza to, że kwadrat przedziału między zdarzeniami w czteroprzestrzeni jest dodatni, a więc sam przedział jest rzeczywisty. Wielkość ta jest niezmiennikiem PL, a więc zachowuje się przy przejściu z jednego układu do drugiego. Dla przedziału takiego mamy

$$ct_{12} > L_{12} \tag{6.27}$$

i nazywamy go **czasopodobnym**. Dwa zdarzenia, które dotyczą tego samego ciała oddzielone są zawsze przedziałem typu czasowego ponieważ poruszające się ciało w czasie t_{12} przejdzie drogę przestrzenną L_{12}, natomiast światło przejdzie drogę ct_{12}. Ponieważ $c > v$, więc $L_{12} < ct_{12}$. Jeżeli $ct_{12} < L_{12}$, wówczas kwadrat odległości jest ujemny i przedział taki nazywamy przedziałem **typu przestrzennego**, natomiast jeżeli $ct_{12} = L_{12}$, to jest to przedział **świetlny**. Przedziałem takiego typu oddzielone są zdarzenia związane z rozchodzeniem się fali świetlnej.

Jeżeli przedział jest typu czasowego, to pojęcia „wcześniej" i „później" mają sens. Dla przedziału typu przestrzennego nie mają, ponieważ

$$c\,dt' = \gamma\left(c\,dt - \frac{v}{c}\,dx\right) = \gamma c\,dt\left(1 - \frac{v}{c}\,\frac{dx}{c\,dt}\right),$$

$$\frac{dt'}{dt} = \gamma\left(1 - \frac{v}{c}\,\frac{dx}{c\,dt}\right).$$

Przyjmijmy, że $dt \geq 0$. Wówczas jeżeli

(1) przedział jest typu czasowego, to $c\,dt > dx$, czyli $dx/(c\,dt) < 1$, a wtedy na pewno

$$1 - \frac{v}{c}\,\frac{dx}{c\,dt} > 0\,,$$

a więc znaki dt i dt' będą takie same;

(2) przedział jest typu przestrzennego, to znaczy $c\,dt < dx$, a wtedy

$$1 - \frac{v}{c}\,\frac{dx}{c\,dt}$$

może być ujemne, czyli znaki dt i dt' mogą być różne.

Jeżeli dwa zdarzenia oddzielone są przedziałem typu czasowego, to może między nimi istnieć powiązanie przyczynowe. Mamy bowiem wtedy

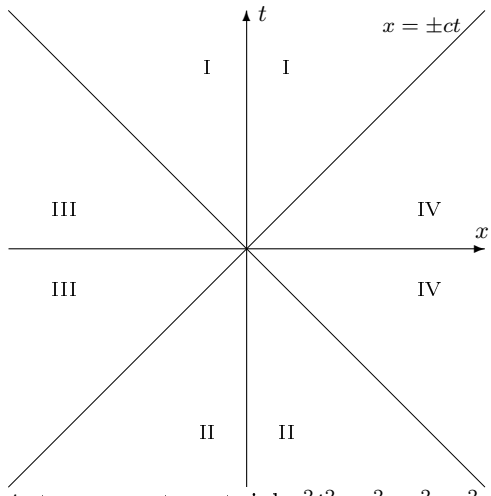
$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 < c^2(t_2 - t_1)^2\,,$$

a więc w ciągu czasu jaki upłynął między zdarzeniem I i II, sygnał świetlny na pewno zdążyłby pokonać odległość przestrzenną między I i II. Mogłaby w ten sposób zostać przekazana informacja do zdarzenia II o zdarzeniu I i mogłoby ono w ten sposób stać się przyczyną zdarzenia II.

Dla przedziałów przestrzennopodobnych kolejność zdarzeń nie jest zachowana i pojęcia „wcześniej" i „później" są względne. Zawsze można znaleźć taki układ w którym kolejność ta będzie odwrócona. Nie istnieje więc związek przyczynowy między zdarzeniami, ponieważ $L_{12} > ct_{12}$, więc w czasie oddzielającym te dwa zdarzenia sygnał świetlny nie zdąży przebyć drogi L_{12}. Nie może być więc przekazana informacja o zdarzeniu I do punktu w którym zajdzie zdarzenie II przed jego nastąpieniem.

Rozpatrzmy teraz dowolne zdarzenie i przyjmijmy je za początek układu współrzędnych czasoprzestrzeni. Dla prostoty przyjmijmy, że zagadnienie ma tylko jedną współrzędną przestrzenną, np x i czas t. W płaszczyźnie (x, t) utwórzmy kartezjański układ współrzędnych z początkiem w zdarzeniu A. Wszystkie zdarzenia będą teraz przedstawiane jako punkty w płaszczyźnie (x, t). Linia świata to tor PM w czteroprzestrzeni. Jednostajny, prostoliniowy ruch PM znajdującego się w chwili $t = 0$ w $x = 0$ to prosta przechodząca przez początek układu. Tangens nachylenia tej prostej to prędkość PM, a jej maksymalna wartość to c. Stąd też linie świata sygnałów świetlnych mają równania $x = \pm ct$. Wszystkie zdarzenia leżące na prostych $x = \pm ct$ oddzielone są przedziałami typu zerowego. Ponieważ największa możliwa prędkość to c, więc linie świata PM mogą leżeć tylko w obszarach I i II.

W obszarach tych przedział między zdarzeniem A i dowolnym innym jest $s_{12} = c^2t^2 - x^2 > 0$, a więc jest czasopodobny, zaś w obszarach III i IV jest przestrzenny. W obszarze I mamy $t > 0$, tzn wszystkie zdarzenia są tam późniejsze niż zdarzenie A. Stąd też, oraz ponieważ dla zdarzeń oddzielonych przedziałem czasowym pojęcia „wcześniej" i „później" są bezwzględne, obszar I nazywamy **absolutną przyszłością** zdarzenia A, natomiast obszar II, z analogicznych powodów, **absolutną przeszłością** zdarzenia A, Zdarzenia leżące w obszarach III i IV oddzielone są od A przedziałami typu przestrzennego, a więc nie może być żadnych związków przyczynowych między nimi a zdarzeniem A.



Powracając do czteroprzestrzeni, zamiast prostych $x = \pm ct$ otrzymamy tam stożek $c^2t^2 - x^2 - y^2 - z^2 = 0$. Stożek ten nazywamy **stożkiem swiatła** . Obszar wewnątrz stożka odpowiada zdarzeniom oddzielonym przedziałami czasowymi od zdarzenia A, natomiast na zewnątrz stożka leżą zdarzenia oddzielone przedziałami przestrzennymi od A.

5. Czterowektory i równania ruchu

Znamy już przekształcenie, które w mechanice relatywistycznej zastąpiło przekształcenie Galileusza oraz znamy przestrzeń w jakiej się ona powinna „rozgrywać", czyli przestrzeń Minkowskiego. Zdarzenia są tam reprezentowane przez punkty, natomiast procesy przez linie świata. Należy teraz znaleźć w tej przestrzeni odpowiedniki wielkości z jakich zbudowana jest mechanika newtonowska. Kierować się będziemy intuicją, tym bardziej, że szukane wielkości powinny, dla małych prędkości, przejść w znane nam wzory mechaniki newtonowskiej. Ściślej, powinny w nie przejść składowe przestrzenne czterowektorów.

Każde zdarzenie można w czasoprzestrzeni określić podając jego czterowektor wodzący x_μ. W mechanice newtonowskiej ruch opisany był przez podanie zależności trzech współrzędnych przestrzennych od czasu, który był niezmiennikiem przekształcenia Galileusza. W mechanice relatywistycznej czas jest traktowany na równi ze współrzędnymi przestrzennymi i nie jest niezmiennikiem PL. Należy więc znaleźć inny parametr i opisać zmianę wszystkich czterech współrzędnych w funkcji tego parametru. Powinien on mieć też charakter czasu i być niezmiennikiem PL, chodzi bowiem o to aby dla małych prędkości powrócić do wzorów mechaniki newtonowskiej. Możemy więc jako parametr wybrać czas własny $d\tau$. Jest on niezmiennikiem PL i dla $c \to \infty$ przechodzi w czas klasyczny - newtonowski. Wektor wodzący x_μ jest przykładem czterowektora, natomiast $d\tau$ to czteroskalar, drugim czteroskalarem jest s_{12} . Znajomość tego

czy dana wielkość jest czteroskalarem, czterowektorem itd. przyda się niedługo przy poszukiwaniu relatywistycznych odpowiedników równań ruchu. Mają one być niezmiennicze wobec PL i ich część przestrzenna powinna dla $c \to \infty$ przechodzić w równania Newtona. Prawa fizyki nie mogą zależeć od tego jak zorientujemy osie układu współrzędnych, muszą więc być niezmiennicze wobec przekształceń ortogonalnych, a w przypadku przestrzeni pseudoeuklidesowej wobec przekształceń pseudoortogonalnych. Osiągane jest to dzięki temu, że przyrównujemy do siebie wielkości tego samego typu, tzn dwa skalary, dwa wektory itd. czyli tensory tego samego rzędu. O równaniu mówi się wówczas, że jest współzmiennicze (kowariantne), ponieważ obie jego strony transformują się w ten sam sposób. Szukamy więc praw fizyki, które byłyby współzmiennicze wobec PL, podobnie jak prawa Newtona były współzmiennicze wobec przekształcenia Galileusza. Jest to więc żądanie aby obie strony równania były czterotensorami tego samego rzędu.

Analogicznie do podziału przedziałów na czaso-, przestrzenno- i światłopodobne możemy podzielić czterowektory. Czterowektor x_μ nazywamy czasopodobnym jeżeli

$$(x,x) = \sum g^{\mu\nu} x_\mu x_\nu > 0\,,$$

przestrzennopodobnym jeżeli

$$(x,x) < 0$$

i świetlnym jeżeli

$$(x,x) = 0\,.$$

Podział ten, podobnie jak dla przedziałów, jest niezmienniczy wobec PL ponieważ (x,x) jest niezmiennikiem PL.

Przystąpmy teraz do skonstruowania czterowektora prędkości. Wiemy, że ruch w czteroprzestrzeni opisywany jest przez linię świata, jest więc to zmiana czterech składowych jako funkcja parametru niezmienniczego wobec PL. Ruch ten powinien być opisywany zależnością $x_\mu = x_\mu(\tau)$. Czterowektor prędkości można określić jako

$$u_\mu = \frac{dx_\mu}{d\tau} = \gamma \frac{dx_\mu}{dt}\,. \tag{6.28}$$

Składowe przestrzenne mają postać

$$u_i = \gamma \frac{dx_i}{dt} = \gamma v_i$$

i zgodnie z naszym żądaniem dla $c \to \infty$ $(\gamma \to 1)$ przechodzą w składowe klasycznej, newtonowskiej prędkości. Składowa czasowa

$$u_0 = \frac{dx_0}{d\tau} = \gamma \frac{d(ct)}{dt} = \gamma c\,.$$

Widać stąd, że jeżeli PM spoczywa, wówczas $v_i = 0$ i składowe przestrzenne czteroprędkości znikają, natomiast $u_0 \neq 0$. Można interpretować to tak, że ruch PM w przestrzeni może być zatrzymany, natomiast w czasie nie.

Czterowektor prędkości możemy zapisać jako $u_\mu = (\gamma c, \gamma \vec{v})$. Sprawdźmy jakiego typu czterowektorem jest czteroprędkość

$$(u,u) = \sum g^{\mu\nu} u_\mu u_\nu = \gamma(c^2 - v^2) = c^2 > 0\,.$$

A więc czteroprędkość jest wektorem typu czasowego. Jest to zresztą oczywiste ponieważ $c > v$.

Policzmy teraz $d(u,u)/d\tau$. Ponieważ $(u,u) = c^2$ oraz iloczyn skalarny jest symetryczny, tzn $(u,v) = (v,u)$, więc $(u, du/d\tau) = 0$. Oznacza to, że czterowektory prędkości, u_μ i przyspieszenia, $du_\mu/d\tau$, są pseudoortogonalne.

Postarajmy się teraz znaleźć relatywistyczne odpowiedniki pędu i siły, tak aby móc napisać relatywistyczne równania ruchu. Newtonowskie prawo ruchu, gdzie masa cząstki była skalarem, ma postać

$$\frac{d}{dt}(m\vec{v}) = \vec{F}\,. \tag{6.29}$$

Mnożąc to skalarnie przez $\vec{v}dt$, otrzymujemy

$$d\left(\frac{mv^2}{2}\right) = \vec{F}\,\vec{v}\,dt\,. \tag{6.30}$$

Jak wiemy jest to zasada zachowania energii cząstki swobodnej. Wyrażenie po lewej stronie to zmiana energii odpowiadająca pracy określonej przez wyrażenie po prawej stronie. Dla cząstki swobodnej możemy więc powiedzieć, że klasyczna energia ciała wyznaczona jest stąd jako

$$E = \frac{1}{2}mv^2 + const.$$

Chcąc aby energia spoczywającego ciała była równa zero, przyjmujemy stałą też za równą zero. Całkowita energia swobodnego, tzn bez sił zewnętrznych, PM jest więc równa jego energii kinetycznej. Formalne uogólnienie drugiego prawa Newtona jest łatwe, zapiszmy je bowiem w postaci

$$F_i = \frac{d}{dt}(mv_i) = \frac{dp_i}{dt}\,.$$

Jego czterowymiarowy odpowiednik powinien mieć więc postać

$$K_\mu = \frac{dp_\mu}{d\tau} \tag{6.31}$$

gdzie K_μ jest czterowektorem siły, tzw **siłą Minkowskiego**, zaś p_μ - **czteropędem**. W analogii do zwykłego pędu możemy p_μ wprowadzić jako

$$p_\mu = mu_\mu\,. \tag{6.32}$$

Ponieważ dla $c \to \infty$ składowe przestrzenne czteroprędkości przechodzą w prędkość newtonowską, więc i składowe przestrzenne pędu przejdą w odpowiednie składowe zwykłego pędu. Mamy dla nich

$$p_i = mu_i = \gamma m v_i\,.$$

Składowe przestrzenne relatywistycznego równania Newtona możemy napisać jako

$$K_i = \frac{dm\,u_i}{d\tau} = \gamma\frac{dm\,u_i}{dt} = \gamma^2\frac{dm\,v_i}{dt}.$$

Jeżeli, znów przez analogię, przyjmiemy $K_i = \gamma F_i$, gdzie F_i są składowymi newtonowskiej siły, to dostaniemy

$$\gamma F_i = m\gamma\frac{du_i}{dt} \quad \Rightarrow \quad F_i = \frac{d}{dt}(mu_i)\,,$$

co jest prędkością zmian relatywistycznego pędu. Nasze relatywistyczne równanie ruchu (6.31) pomnóżmy skalarnie przez u_μ

$$(K_\mu, u_\mu) = (\frac{dp_\mu}{d\tau}, u_\mu) = m(\frac{du_\mu}{d\tau}, u_\mu) = 0\,,$$

ponieważ czteroprędkość jest ortogonalna do czteroprzyspieszenia. Stąd

$$K_0u_0 - \sum K_iu_i = 0 \quad \Rightarrow \quad K_0 = \frac{\sum K_iu_i}{c\gamma} = \frac{\gamma}{c}\sum F_iv_i\,.$$

Tak więc relatywistyczne równania Newtona można zapisać oddzielnie dla części przestrzennej

$$K_i = \frac{d}{d\tau}mu_i \quad \text{albo} \quad F_i = \frac{d}{dt}mu_i \tag{6.33}$$

i dla części czasowej

$$K_0 = \frac{d}{d\tau} m u_0 \quad \text{albo} \quad \frac{\gamma}{c} \vec{F}\,\vec{v} = \gamma \frac{d m u_0}{dt}\,. \tag{6.34}$$

lub

$$d(mc^2\gamma) = \vec{F}\,\vec{v}\,dt\,. \tag{6.35}$$

Odpowiednikiem ostatniego równania w mechanice newtonowskiej jest

$$d\left(\frac{mv^2}{2}\right) = \vec{F}\,\vec{v}\,dt \quad \Rightarrow \quad E = \frac{mv^2}{2}\,.$$

W mechanice relatywistycznej ostatnia równość ma postać

$$E = \gamma mc^2 + const. \tag{6.36}$$

Wybieramy $const = 0$.

Zobaczmy teraz jaki jest sens składowej czasowej czteropędu. Jak wynika z (6.32)

$$p_0 = m u_0 = \gamma m c = \frac{E}{c}\,. \tag{6.37}$$

Składowa ta jest więc proporcjonalna do energii. Stąd też często mówi się o **czterowektorze energii-pędu**. Możemy go zapisać jako

$$p_\mu = \left(\frac{E}{c}, p_i = \gamma m v_i\right)\,. \tag{6.38}$$

p_i jest relatywistycznym pędem trójwymiarowym.

Równanie (6.36) jest całkowitą energią swobodnej cząstki, która dla $c \to \infty$ przechodzi we wzór klasyczny. Rozwińmy bowiem w szereg wyrażenie $E = \gamma mc^2$ dla małych (w porównaniu z c) prędkości

$$E = \frac{mc^2}{\sqrt{1-\left(\frac{v}{c}\right)^2}} = mc^2\left[1 + \frac{1}{2}\left(\frac{v}{c}\right)^2 + \frac{3}{8}\left(\frac{v}{c}\right)^4 - \ldots\right] \simeq$$

$$\simeq mc^2 + \frac{mv^2}{2} + \frac{3}{8} m \frac{v^4}{c^2} - \ldots \quad \simeq mc^2 + \frac{mv^2}{2}$$

Jeżeli cząstka spoczywa ($\vec{v} = 0$), wówczas jej energia (tzw energia spoczynkowa) jest równa

$$E_{spocz} = mc^2\,. \tag{6.39}$$

Masę m nazywamy masą spoczynkową. Energia (6.39) jest stała, moglibyśmy więc tak dobrać stałą występującą w definicji energii (6.36) aby upraszczała się z E_{spocz}, złamałoby to jednak korespondencję z przypadkiem nierelatywistycznym. Wykorzystajmy teraz PL dla dowolnego czterowektora A_μ, która dla przypadku gdy układy poruszają się wzdłuż osi Ox ma postać

$$A_0{}' = \gamma(A_0 - \beta A_1), \quad A_1{}' = \gamma(A_1 + \beta A_0). \tag{6.40}$$

Dla składowych przestrzennych czteropędu będziemy mieli

$$p_x' = \gamma\left(p_x - \frac{v}{c^2}E\right),$$

gdzie v to względna prędkość obu układów.

$$m v_x' = \gamma\left(m v_x - \frac{v}{c^2}E\right).$$

Przyjmijmy teraz, że

$$E = \gamma mc^2 + const = \gamma mc^2 - mc^2$$

tak aby stała upraszczała się z energią spoczynkową. Mamy wtedy

$$mv'_x = \gamma m v_x - \gamma^2 v m - \gamma v \frac{(-mc^2)}{c^2} = \gamma m v_x \,.$$

Po przejściu do granicy nierelatywistycznej ($\gamma \to 1$) i podzieleniu przez m otrzymalibyśmy

$$v'_x = v_x \,,$$

a więc sprzeczność z transformacją Galileusza (niezależność od względnej prędkości układów).

Możemy wprowadzić nową wielkość - **masę relatywistyczną**, zależną od prędkości

$$m_r = \gamma m \tag{6.41}$$

i wtedy czterowektor energii-pędu

$$p_\mu = \left(\frac{E}{c}, p_i = m_r v_i\right).$$

Obliczmy teraz kwadrat czteropędu

$$(p,p) = \left(\frac{E}{c}\right)^2 - p^2 = m^2(u,u) = m^2c^2 > 0\,.$$

Jest to wielkość stała, a sam czterowektor pędu jest wektorem czasowym. Stąd też otrzymujemy relatywistyczny związek między energią i pędem

$$E = c\sqrt{p^2 + m^2c^2}\,. \tag{6.42}$$

Jeżeli rozpatrywany przez nas PM nie jest swobodny lecz znajduje się w polu o potencjale V, wówczas zamiast (6.36) i (6.42) dostaniemy

$$E = \gamma mc^2 + V \quad \text{oraz} \quad E = c\sqrt{p^2 + m^2c^2} + V.$$

Mechanika newtonowska powinna być zastąpiona przez relatywistyczną gdy $v \cong c$. Ten sam warunek możemy też sformułować inaczej, jako

$$\gamma = \left(1 - \left(\frac{v}{c}\right)^2\right)^{-1/2} \gg 1\,,$$

a to jest równoważne warunkowi

$$\gamma = \frac{E}{mc^2} \gg 1\,.$$