

Elementy Fizyki Teoretycznej

Stanisław Bednarek

(semestr letni 2009)

I.

- Wstęp.
- Mechanika klasyczna.
- Wprowadzenie.
- Więzy.
- Współrzędne uogólnione.
- Zasada najmniejszego działania (elementy rachunku wariacyjnego).

II.

- Transformacja Galileusza, jednorodność i izotropowość przestrzeni i czasu.
- Funkcja Lagrange'a cząstki swobodnej w inercjalnym układzie odniesienia.
- Funkcja Lagrange'a dla cząstki w polu potencjalnym.
- Funkcja Lagrange'a układu wielu cząstek.
- Transformacja ze współrzędnych kartezjańskich do współrzędnych uogólnionych.
- Funkcja Lagrange'a dla cząstki swobodnej we współrzędnych sferycznych i cylindrycznych.
- Przykładowe zastosowanie formalizmu Lagrange'a.
- Cząstka w polu centralnego potencjału.
- Wahadło matematyczne.

III.

- Bardziej zaawansowane zastosowania formalizmu Lagrange'a.
- Wahadło złożone.
- Wydzielenie podukładu z układu cząstek oddziaływujących.
- Problem wahadła matematycznego, którego punkt zawieszenia wykonuje pionowe drgania.
- Równania ruchu w nieinercjalnych układach odniesienia.

IV.

- Prawa zachowania, całki ruchu.
- Energia.
- Pęd.
- Moment pędu.
- Zmienne cykliczne a całki ruchu.
- Równania kanoniczne Hamiltona.
- Nawiasy Poissona.

(2) Mechanika relatywistyczna

V.

- Zasada względności Einsteina
- Przedział czasoprzestrzenny.
- Transformacja Lorentza.

- Transformacja odwrotna.
- Kontrakcja długości i dylatacja czasu.
- Transformacja prędkości.

VI.

- Relatywistyczna całka działania.
- Pęd i energia cząstki swobodnej.
- Relatywistyczna zasada zachowania energii.
- Relatywistyczna funkcja Hamiltona.
- Transformacja energii i pędu, czterowektory.

VII.

- Geometria czasoprzestrzeni, elementy rachunku tensorowego.
- Tensory kontrawariantne i kowariantne.
- Tensor metryczny.
- Tensorowe własności operatorów różniczkowych.

(3) Elektrodynamika

VIII.

- Cząstka w zewnętrznym polu elektrycznym i magnetycznym, czteropotencjał pola elektromagnetycznego.
- Interpretacja czteropotencjału.
- Niezmienniczość cechowania potencjałów elektromagnetycznych.
- Transformacja Lorentza dla pól elektrycznego i magnetycznego.

IX.

- Pierwsza para równań Maxwella.
- Czterowektor gęstości prądu, równanie ciągłości.
- Pole elektromagnetyczne wywołane zadaniem rozkładem ładunków i prądów.
- Działanie pola elektromagnetycznego, druga para równań Maxwella.
- Równania Maxwella w postaci różniczkowej.
- Kowariantny (jawnie relatywistyczny) zapis równań Maxwella.

X.

- Równania Maxwella w różniczkowej i całkowej postaci.
- Refleksja nad sposobem wprowadzenia równań Maxwella.
- Zastosowanie równań Maxwella w postaci całkowej.
- Prawo Coulomba.
- Potencjał skalarny ładunku punktowego.
- Układ kilku ładunków punktowych.
- Ciągły rozkład gęstości ładunku.
- Pole elektryczne wokół jednorodnie naładowanego walca.
- Symetria prostokątna- jednorodnie naładowana płaszczyzna, kondensator.
- Zastosowania praw Maxwella w postaci całkowej do magnetostatyki.
- Pole magnetyczne wewnątrz nieskończenie długiego solenoidu.

XI.

- Transformacja Lorentza w magnetostatyce.
- Zastosowania równań Maxwella w postaci różniczkowej.
- Równanie Poissona i Laplace'a.
- Twierdzenie o jednoznaczności rozwiązań równania Laplace'a i Poissona.
- Prawo Biota-Savarta.

XII.

- Przewodniki i warunki brzegowe na ich powierzchniach.
- Ładunki indukowane.
- Pole elektrostatyczne w obecności przewodników. Metoda obrazów.
- Obraz a ładunek indukowany.
- Rozwinięcie multipolowe

XIII.

- Pola elektryczne i magnetyczne w ośrodkach, zależności pomiędzy E i D oraz H i B .
- Istota wektorów pól.
- Polaryzacja dielektryka.
- Dielektryki liniowe.
- Granice ośrodków dielektrycznych
- Zmienne pole elektromagnetyczne. Fala płaska.

XIV.

- Elektrodynamika klasyczna w nanostrukturach półprzewodnikowych.

I.

Wstęp.

Fizykę można podzielić w zależności od problemów jakimi się zajmuje wyróżnić różne jej działy, np. fizyka jądrowa, ciała stałego, niskich temperatur, astrofizyka itd.

Niezależnie od podziału tematycznego istnieje podział na fizykę doświadczalną i teoretyczną.

Ta ostatnia wyróżnia się specyficznymi narzędziami pracy. Są nimi kartka papieru, ołówek i komputer jako narzędzia materialne i aparat matematyczny, który możemy określić jako niematerialne narzędzia pracy teoretyka. Celem niniejszego wykładu jest zapoznanie słuchaczy z kilkoma takimi niematerialnymi narzędziami pracy używanymi w kilku wybranych działach fizyki. Wykład trwa tylko jeden semestr, w związku z czym nie jest możliwa ani kompleksowa ani dogłębna analiza problemów. Wybrałem problemy, które według mojej oceny są najważniejsze lub najciekawsze. Podzielić je można na następujące grupy tematyczne:

I. Mechanika klasyczna

II. Mechanika relatywistyczna

III. Elektrodynamika

Łączyć je będzie zasada najmniejszego działania, przy pomocy której uzyskane zostaną wszystkie podstawowe prawa fizyki. Prześledzimy jednocześnie ewolucję funkcji działania. Zostanie ona początkowo zaproponowana dla mechaniki klasycznej następnie uogólniona do mechaniki relatywistycznej, a kolejna modyfikacja pozwoli na objęcie nią elektrodynamiki.

Klasyczna Mechanika Teoretyczna.

Jest to dział Fizyki, który opisuje ruch bardziej lub mniej złożonych układów fizycznych. Na wstępie musimy pogodzić się z faktem, iż nie jesteśmy w stanie opisać całego wszechświata w całej jego złożoności. Jesteśmy zmuszeni posłużyć się przybliżeniami i ograniczyć się do opisu jego niewielkiej części. Przystępując do analizy jakiegokolwiek zjawiska musimy wypunktować wszystkie przybliżenia i założenia wchodzące do rachunku. Ich spełnienie jest warunkiem wiarygodności uzyskanych wyników. W celu uniknięcia nieporozumień językowych ustalić musimy na wstępie kilka najważniejszych pojęć i przybliżeń.

- Układem nazywać będziemy część wszechświata, którą opisujemy. Możemy założyć brak oddziaływania układu z resztą wszechświata, mówimy wówczas o układzie izolowanym (np. cząstka swobodna), lub uwzględniamy to oddziaływanie traktując je jako oddziaływanie zewnętrzne (np. ciało w polu sił ciężkości).
- Najprostszym układem (lub jego elementem) jest punkt materialny (cząstka), który rozumieć będziemy jako obiekt fizyczny, którego rozmiary można zaniedbać przy opisie jego ruchu (np. elektron w polu jądra, planeta na orbicie).
- Przestrzeń, w której rozgrywa się mechanika punktu materialnego jest trójwymiarowa. Położenie punktu materialnego opisujemy trójwymiarowym **wektorem wodzącym** $\vec{r}$.
- Do zapisu wektora wodzącego posługujemy się układem współrzędnych. Najczęściej stosować będziemy układ prostokątny kartezjański, lub krzywoliniowy: sferyczny, cylindryczny, paraboliczny.
- Równie ważną rolę jak przestrzeń, w której rozgrywać się będą zjawiska spełnia czas. Będziemy go rozumieli jako parametr numerujący kolejność zdarzeń.
- Pełną informację o ruchu pojedynczego punktu materialnego zawiera jego trajektoria (tor) czyli położenie w funkcji czasu $\vec{r}(t)$ i stanowi szukane rozwiązanie problemu mechanicznego.
- Znając trajektorię możemy wyliczyć prędkość jako pierwszą i przyspieszenie jako drugą pochodną wektora wodzącego po czasie, a z nich możemy wyliczyć wszystkie pozostałe wielkości fizyczne np. energię, pęd, siły. Uwaga: pochodne po czasie oznaczать będziemy kropką: $\frac{d}{dt} \vec{r}(t) = \dot{\vec{r}}(t), \frac{d^2}{dt^2} \vec{r}(t) = \ddot{\vec{r}}(t)$.
- Związki pomiędzy współrzędnymi, prędkościami i przyspieszeniami nazywamy równaniami ruchu. Są to na ogół równania różniczkowe 2 rzędu, w których niewiadomą funkcję stanowi trajektoria $\vec{r}(t)$ (np. równania Newtona).
- Przy znanych równaniach ruchu warunkiem jednoznaczności rozwiązania jest podanie warunków początkowych. Mogą być nimi położenie i prędkość cząstki w wybranej chwili. Warunki te określają stan układu.
- W przypadku gdy układ składa się z N punktów materialnych (cząstek), rozwiązaniem problemu mechanicznego jest $3N$ -wymiarowa trajektoria czyli podanie wektorów wodzących wszystkich cząstek wchodzących w skład układu.
- Liczbę niezależnych współrzędnych konieczną do opisu układu nazywamy liczbą stopni swobody i oznaczamy literą f . Cząstka swobodna w przestrzeni trójwymiarowej ma 3, a układ N cząstek $3N$ stopni swobody.

Więzy.

Liczbę stopni swobody ograniczyć mogą tzw. więzy, czyli związki pomiędzy współrzędnymi, które muszą być spełnione w dowolnej chwili czasu. Mogą być zadane w postaci równości lub nierówności:

$$f_i(\vec{r}_1, \dots, \vec{r}_N, t) = 0, \quad i=1 \dots n_w$$

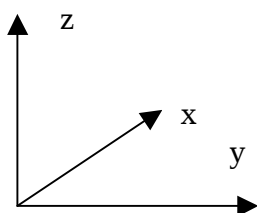
lub

$$f_i(\vec{r}_1, \dots, \vec{r}_N, t) \geq 0, \quad i=1 \dots n_w$$

Pierwszy typ więzów nazywamy dwustronnymi, drugi jednostronnymi. Jeżeli równania więzów nie zawierają jawnej zależności czasowej noszą nazwę skleronomicznych, w przeciwnym wypadku reonomicznych. Więzy narzucone na współrzędne noszą nazwę więzów holonomicznych. Czasem spotykamy się z więzami narzucenymi na prędkości, są one nieholonomiczne. Każde niezależne równanie więzów dwustronnych ogranicza liczbę stopni swobody układu o 1.

Przykłady więzów:

1. Ruch cząstki w płaszczyźnie dwuwymiarowej (x,y) możemy uzyskać narzucając funkcję więzów:



$$f(\vec{r}) = z = 0$$

Liczba więzów $p=1$, liczba stopni swobody $f=2$.

2. Ruch po okręgu :

$$f_1(\vec{r}) = z$$

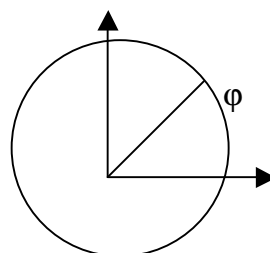
$$f_2(\vec{r}) = x^2 + y^2 - R^2$$

$p=2$, $f=1$.

W obu przykładach podane zostały więzy holonomiczne, skleronomiczne, dwustronne.

3. Ruch piłki na boisku. (więzy jednostronne).

$$z \geq 0$$



Współrzędne uogólnione.

Występowanie więzów powoduje, że w równaniach ruchu uwzględniać musimy dodatkowe siły reakcji więzów. Wiąże się to ze znaczną komplikacją równań. Często komplikacji udaje się uniknąć przez odpowiedni wybór układu współrzędnych. Np. w przykładzie 2 wystarczy przejść do współrzędnych biegunowych. Wówczas więzy ustalają odległość cząstki od początku układu. Współrzedną odpowiadającą jednemu stopniowi swobody jest kąt, który nie jest ograniczany przez więzy.

W ten sposób dochodzimy do korzystnego z punktu widzenia ułatwień rachunkowych pojęcia współrzędnych uogólnionych.

Współrzednymi uogólnionymi układu o f stopniach swobody nazywamy dowolny zbiór f wielkości $q_1, \dots, q_f$ wyznaczających jednoznacznie położenia ciał tworzących układ.

$$q = \{q_\alpha\} (\alpha = 1, \dots, f)$$

Pochodne po czasie współrzędnych uogólnionych nazywamy prędkościami uogólnionymi.
 $\dot{q} = \{\dot{q}_\alpha\} (\alpha = 1, \dots, f)$

np. w układzie biegunowym $q = \{r, \phi\}, \dot{q} = \{\dot{r}, \dot{\phi}\}$. Naturalnie współrzędne i prędkości w układzie kartezjańskim $q = \{x, y, z\}, \dot{q} = \{\dot{x}, \dot{y}, \dot{z}\}$ stanowią również szczególny przypadek współrzędnych i prędkości uogólnionych, o ile liczba stopni swobody nie jest ograniczona przez więzy.

Ze względu na uniwersalność współrzędnych uogólnionych, wszystkie ogólne rozważania w dalszej części wykładu będą prowadzone w tych właśnie współrzędnych. Dopiero w konkretnych zastosowaniach będziemy przechodzili do najwygodniejszych (ze względu na symetrię układu) konkretnych układów współrzędnych.

Zasada najmniejszego działania (elementy rachunku wariacyjnego).

Całą mechanikę klasyczną można oprzeć (historycznie tak ona powstała) na równaniach Newtona i potraktować je jako elementarne, startowe założenie teorii służącej do opisu zjawisk. Jednakże w wielu przypadkach jest to niewygodne, szczególnie w momencie przejścia do „nieklasycznej” teorii.

Przez prawie dwieście lat po Newtonie fizycy poszukiwali uogólnień i bardziej uniwersalnych podejść.

Takim udanym uogólnieniem teorii prowadzącym do praw ruchu w postaci najbardziej odpowiadającej rozważanemu układowi jest zasada najmniejszego działania, lub inaczej zasada wariacyjna Hamiltona.

Rozpatrzmy układ o f stopniach swobody, opisywany współrzędnymi uogólnionymi q_α .

Niech zbiór funkcji:

$$q_\alpha = q_\alpha(t) \text{ gdzie } \alpha = 1, \dots, f$$

opisuje ruch rzeczywisty układu i stanowi rozwiązanie problemu.

Zdefiniujmy inny zbiór funkcji, zmieniony w porównaniu do ruchu rzeczywistego o niewielką wartość w każdej chwili czasowej:

$$\tilde{q}_\alpha = q_\alpha(t) + \delta q_\alpha(t)$$

i nazwijmy go ruchem porównawczym. $\delta q_\alpha(t)$ nieskończenie małe odchylenie od ruchu rzeczywistego nazywamy wariacją współrzędnej uogólnionej.

Jeżeli przez $F(q_\alpha, \dot{q}_\alpha, t)$ oznaczymy dowolną funkcję współrzędnych i prędkości uogólnionych oraz czasu, możemy wyliczyć jej wariację jako różnicę pomiędzy funkcją od argumentów odpowiadających ruchowi porównawczemu i rzeczywistemu:

$$\delta F = F(\tilde{q}_\alpha, \tilde{\dot{q}}_\alpha, t) - F(q_\alpha, \dot{q}_\alpha, t) = F(q_\alpha + \delta q_\alpha, \dot{q}_\alpha + \delta \dot{q}_\alpha, t) - F(q_\alpha, \dot{q}_\alpha, t) =$$

$$= \sum_{\alpha} \left(\frac{\partial F}{\partial q_\alpha} \delta q_\alpha + \frac{\partial F}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha \right)$$

gdzie $\delta \dot{q}_\alpha(t) = \tilde{\dot{q}}_\alpha - \dot{q}_\alpha$ stanowi wariację prędkości uogólnionej. Oczywiście ze względu na liniowość:

$$\delta \dot{q}_\alpha(t) = \frac{d}{dt} \delta q_\alpha(t)$$

Możemy również zdefiniować wariację funkcjonału. Rozważmy funkcjonał w postaci:

$$I[q_\alpha, \dot{q}_\alpha] = \int_{t_1}^{t_2} dt F(q_\alpha, \dot{q}_\alpha, t)$$

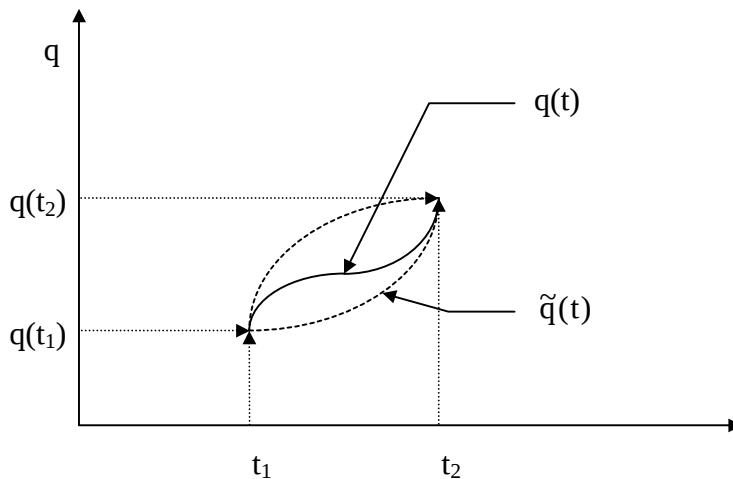
Jego wariacja

$$\delta I = I[\tilde{q}_\alpha, \tilde{\dot{q}}_\alpha] - I[q_\alpha, \dot{q}_\alpha] = \int_{t_1}^{t_2} dt (F(\tilde{q}_\alpha, \tilde{\dot{q}}_\alpha, t) - F(q_\alpha, \dot{q}_\alpha, t)) = \int_{t_1}^{t_2} dt \delta F$$

Możemy teraz sformułować zasadę najmniejszego działania.

Rozważamy ruch rzeczywisty i porównawcze w pewnym określonym przedziale czasu (t_1, t_2) , przy założeniu, że w chwilach odpowiadających początkowi i końcowi przedziału współrzędne uogólnione wszystkich ruchów przyjmują jednakowe ustalone wartości:

$$\tilde{q}_\alpha(t_1) = q_\alpha(t_1) \text{ oraz } \tilde{q}_\alpha(t_2) = q_\alpha(t_2) \text{ czyli } \delta q_\alpha(t_1) = \delta q_\alpha(t_2) = 0$$



Postulat:

Dla każdego układu mechanicznego można znaleźć funkcję współrzędnych i prędkości uogólnionych oraz czasu $L(q_\alpha, \dot{q}_\alpha, t)$ taką, że funkcjonał:

$$S[\tilde{q}_\alpha, \tilde{\dot{q}}_\alpha] = \int_{t_1}^{t_2} dt L(\tilde{q}_\alpha, \tilde{\dot{q}}_\alpha, t)$$

przyjmuje najmniejszą wartość dla ruchu rzeczywistego, tj. $\tilde{q}_\alpha(t) = q_\alpha(t)$. Funkcja L nosi nazwę funkcji Lagrange'a, a funkcjonał S nazywamy działaniem.

Warunek najmniejszego działania dla ruchu rzeczywistego generuje równania ruchu we współrzędnych uogólnionych.

Narzućmy ten warunek, czyli zażądajmy $\delta S = 0$.

$$\begin{aligned} \delta S &= \delta \int_{t_1}^{t_2} dt L(q_\alpha, \dot{q}_\alpha, t) = \int_{t_1}^{t_2} dt \delta L(q_\alpha, \dot{q}_\alpha, t) = \int_{t_1}^{t_2} dt \sum_\alpha \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha \right) = \\ &= \sum_\alpha \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha + \frac{\partial L}{\partial \dot{q}_\alpha} \frac{d}{dt} \delta q_\alpha \right) \\ &= \sum_\alpha \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \frac{d}{dt} \delta q_\alpha \right) + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \delta q_\alpha \right) \\ &= \sum_\alpha \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_\alpha} \delta q_\alpha - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) \delta q_\alpha \right) + \sum_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} \delta q_\alpha \Big|_{t_1}^{t_2} = \sum_\alpha \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} \right) \delta q_\alpha = 0 \end{aligned}$$

Ponieważ warunek ten musi być spełniony dla dowolnego wyboru wariacji współrzędnych uogólnionych, musi się zerować wyrażenie w nawiasie:

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = 0$$

Równania te noszą nazwę równań Lagrange'a (drugiego rodzaju). Stanowią one układ równań ruchu (f równań różniczkowych drugiego rzędu) na funkcje $q_\alpha(t)$, stanowiące trajektorie ruchu rzeczywistego we współrzędnych uogólnionych.

Zanim skonstruujemy funkcję Lagrange'a dla najprostszego układu zwróćmy uwagę na pewne jej własności wynikające z jej definicji i uzyskanych przy jej pomocy równań ruchu..

- Funkcja Lagrange'a zapisana w różnych układach współrzędnych uogólnionych może prowadzić do różnych związków pomiędzy nimi (równań ruchu).
- Jeżeli funkcja Lagrange'a zapisana w jednych współrzędnych prowadzi do poprawnych równań ruchu, wówczas równania ruchu uzyskane z tej samej funkcji Lagrange'a zapisanej w innych współrzędnych są również poprawne (przetransformowane). Transformacja funkcji Lagrange'a często jest łatwiejsza niż transformacja gotowych równań ruchu.
- Funkcja Lagrange'a układu złożonego z dwóch nie oddziałujących podukładów jest sumą funkcji Lagrange'a obu części.

$$L_{AB} = L_A + L_B$$

- Funkcja Lagrange'a nie jest zdefiniowana jednoznacznie.
- * Pomnożenie funkcji Lagrange'a przez dowolną stałą (różną od zera) prowadzi do tych samych równań ruchu.

$$L' = aL \Leftrightarrow L$$

- * Dodanie do funkcji Lagrange'a pochodnej zupełnej po czasie dowolnej funkcji współrzędnych i prędkości uogólnionych prowadzi do tych samych równań ruchu.

$$L' = L + \frac{d}{dt} f(q_\alpha, \dot{q}_\alpha, t) \Leftrightarrow L$$

Własności powyższe są dość oczywiste, pierwsza wynika z liniowości równań Lagrange'a, druga daje to samo działanie, gdyż po scałkowaniu i narzuceniu warunków początkowych pochodna zupełna dowolnej funkcji daje zero. Zwróćmy jednak uwagę na odróżnienie pochodnej po czasie cząstkowej od zupełnej. Dowolna funkcja współrzędnych, prędkości i czasu $f(q, \dot{q}, t)$ dla ruchu rzeczywistego zawiera dodatkową zależność czasową poprzez współrzędne i prędkości. Jej różniczka zupełna ma postać:

$$df(q, \dot{q}, t) = \frac{\partial f}{\partial q} dq + \frac{\partial f}{\partial \dot{q}} d\dot{q} + dt,$$

a zupełna pochodna po czasie:

$$\frac{d}{dt} f(q, \dot{q}, t) = \frac{\partial f}{\partial q} \frac{dq}{dt} + \frac{\partial f}{\partial \dot{q}} \frac{d\dot{q}}{dt} + \frac{\partial f}{\partial t} = \frac{\partial f}{\partial q} \dot{q} + \frac{\partial f}{\partial \dot{q}} \ddot{q} + \frac{\partial f}{\partial t}$$

Podsumowanie I.

1. We wprowadzeniu ustaliliśmy czym zajmować się będziemy podczas niniejszego wykładu.
2. Podane zostały podstawowe pojęcia mechaniki klasycznej: układ, punkt materialny, przestrzeń, czas, trajektoria, równanie ruchu, stan układu, więzy, stopnie swobody.
3. Więzy dwustronne ograniczają liczbę stopni swobody, zazwyczaj każde równanie więzów obniża liczbę stopni swobody o 1.
4. Wprowadziliśmy pojęcie współrzędnych uogólnionych, które stanowią zbiór wielkości, wyznaczających jednoznacznie położenie ciała. Ich liczba jest równa liczbie stopni swobody.
5. Wprowadziliśmy elementy rachunku wariacyjnego definiując trajektorie ruchu rzeczywistego i porównawczych, różniące się o małą wielkość, którą nazwaliśmy wariacją współrzędnej uogólnionej.
6. Wprowadziliśmy pojęcie wariacji funkcji współrzędnych i prędkości uogólnionych oraz wariacji utworzonego z takich funkcji funkcjonału.
7. Zapostulowaliśmy zasadę najmniejszego działania, zgodnie z którą funkcjonał:

$$S[\tilde{q}_\alpha, \dot{\tilde{q}}_\alpha] = \int_{t_1}^{t_2} dt L(\tilde{q}_\alpha, \dot{\tilde{q}}_\alpha, t)$$

osiąga minimum dla ruchu rzeczywistego.

8. Żądanie znikania wariacji działania prowadzi do równań Lagrange'a, które powinny stanowić równania ruchu dla współrzędnych uogólnionych:

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = 0$$

9. Funkcji Lagrange'a L jeszcze nie umiemy konstruować, ale poznaliśmy już jej pewne własności wynikające ze sposobu jej wprowadzenia.

II.

Transformacja Galileusza, jednorodność i izotropowość przestrzeni i czasu.

Konstrukcja funkcji Lagrange'a bazuje na podstawowych własnościach przestrzeni i czasu, które muszą być uwzględnione przy tworzeniu równań ruchu. Równania ruchu oczywiście zależą od wyboru układu współrzędnych i nieco szerzej rozumianego układu odniesienia. Podstawowymi układami odniesienia, w których będziemy pracować będą tzw. układy inercjalne.

Inercjalnym układem odniesienia nazywamy układ odniesienia, w którym ruch swobodny odbywa się ze stałą prędkością (np. dla cząstki w polu sił ciężkości układ związany z powierzchnią ziemi nie jest inercjalny, ale inercjalny jest układ związany ze swobodnie spadającą windą).

Własność inercjalności jest równoważna jednorodności i izotropowości przestrzeni oraz jednorodności czasu. Oznacza to, że w układzie inercjalnym nie ma wyróżnionych położeń, kierunków i chwil.

Sformułujmy teraz bardzo ważne założenie teorii:

Wszystkie układy odniesienia poruszające się względem pewnego inercjalnego układu odniesienia ze stałą prędkością są również inercjalne.

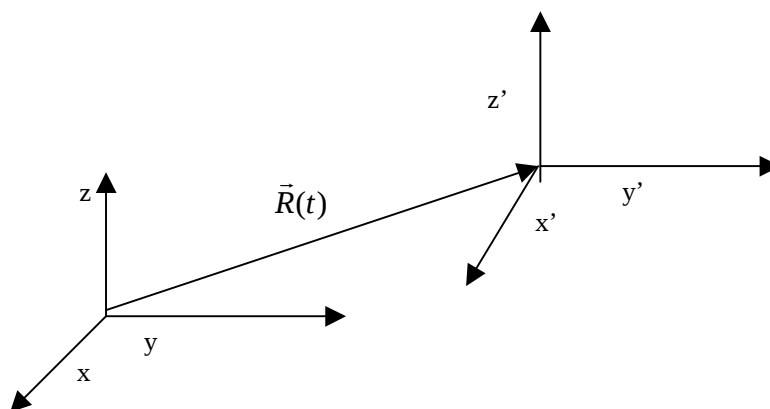
Założenie to wymusza jednolity opis zjawisk fizycznych we wszystkich układach odniesienia poruszających się względem siebie z jednakowymi prędkościami a jednocześnie pozwala na ich swobodny wybór.

Związek pomiędzy położeniami i prędkościami przy przejściu pomiędzy dwoma inercjalnymi układami odniesienia (słuszny w mechanice klasycznej) nosi nazwę Transformacji Galileusza.

Jeżeli mamy dwa układy odniesienia (x, y, z) i (x', y', z') , przy czym układ "primowany"

porusza się względem "nie primowanego" z pewną stałą prędkością $\vec{V}$, a więc położenie początku układu "primowanego" dane jest w układzie "nie primowanym" związkiem:

$$\vec{R}(t) = \vec{R}(0) + \vec{V}t$$



wówczas punkt o współrzędnych $\vec{r}'$ ma w układzie "nie primowanym" współrzędne

$$\vec{r} = \vec{R}(t) + \vec{r}' = \vec{R}(0) + \vec{V}t + \vec{r}'$$

Pomiędzy prędkościami mamy związek:

$$\vec{v} = \vec{V} + \vec{v}'$$

Związek pomiędzy prędkościami wynika z założenia czasu bezwzględnego, czyli jednakowego we wszystkich inercjalnych układach odniesienia (w mechanice relatywistycznej nie dokonuje się tego założenia).

Funkcja Lagrange'a cząstki swobodnej w inercjalnym układzie odniesienia.

Założmy, że układ, który opisujemy składa się z pojedynczej nie oddziaływującej z otoczeniem cząstki. Układ posiada trzy stopnie swobody, położenie cząstki opisuje wektor wodzący $\vec{r}$, a prędkość jego pochodna po czasie $\vec{v} = \dot{\vec{r}}$. Ich składowe potraktujemy jako współrzędne i prędkości uogólnione. Będą więc one stanowić argumenty funkcji Lagrange'a $L(q, \dot{q}, t) = L(\vec{r}, \vec{v}, t)$

Ze względu na jednorodność przestrzeni (brak wyróżnionego punktu) funkcja Lagrange'a nie powinna zależeć od położenia a ze względu na jednorodność czasu nie powinna również od niego zależeć. Jedynym argumentem, od którego może funkcja Lagrange'a zależeć jest prędkość. Ale przestrzeń jest izotropowa. Brak wyróżnionego kierunku w przestrzeni pozostawia jedynie zależność od bezwzględnej wartości prędkości. Najprostszą taką funkcją jest :

$$L = \alpha v^2.$$

Przyjęcie funkcji Lagrange'a w tej postaci zapewnia jednocześnie niezmienniczość równań ruchu przy przejściu pomiędzy różnymi układami inercjalnymi. Zauważmy, że

$$L(\vec{v}) = \alpha \vec{v}^2 = \alpha (\vec{v}' + \vec{V})^2 = \alpha \vec{v}'^2 + 2\alpha \vec{v}' \cdot \vec{V} + \alpha \vec{V}^2 = \alpha \vec{v}'^2 + \frac{d}{dt}(2\alpha \vec{r}' \cdot \vec{V} + \alpha \vec{V}^2 t) \Leftrightarrow$$

$$\Leftrightarrow L(\vec{v}') = \alpha \vec{v}'^2$$

Założenie innej, np. nieliniowej zależności funkcji Lagrange'a od kwadratu prędkości jest dopuszczalne ze względu na jednorodność i izotropowość przestrzeni i czasu, nie daje jednak niezmienniczości równań ruchu przy przejściach pomiędzy inercjalnymi układami odniesienia.

W uzyskanym wyrażeniu stała α nie jest wyznaczona jednoznacznie i może być przyjęta dowolnie. Jeżeli przyjmujemy $\alpha = \frac{m}{2}$, uzyskamy:

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

gdzie przez T oznaczamy energię kinetyczną układu.

Wykorzystajmy uzyskaną postać funkcji L i wstawmy ją do równań Lagrange'a:

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = 0$$

Otrzymujemy:

$$0 = \frac{\partial}{\partial \vec{r}_i} \left(\frac{m}{2} \dot{\vec{r}}^2 \right) - \frac{d}{dt} \frac{\partial}{\partial \dot{\vec{r}}_i} \left(\frac{m}{2} \dot{\vec{r}}^2 \right) = -\frac{d}{dt} m \dot{\vec{r}}_i = -m \ddot{\vec{r}}_i$$

czyli

$$m \ddot{\vec{r}} = 0$$

jest to znane nam równanie Newtona dla ruchu cząstki swobodnej w inercjalnym układzie odniesienia.

Zauważmy, że z zupełnie ogólnych rozważań uzyskaliśmy precyzyjnie wyznaczoną (z dokładnością do będących jej ogólną własnością niejednoznaczności) funkcję Lagrange'a dla cząstki swobodnej a uzyskane z równań Lagrange'a równanie ruchu jest poprawne. Postulat spełnienia zasady najmniejszego działania jest więc dla cząstki swobodnej słuszny. Uzyskany wynik możemy natychmiast uogólnić na układ N nie oddziaływujących z otoczeniem i pomiędzy sobą cząstek:

$$L = \sum_{a=1}^N L_a = \sum_{a=1}^N T_a = T$$

Funkcję Lagrange'a takiego układu stanowi jego całkowita energia kinetyczna.

Funkcja Lagrange'a dla cząstki w polu potencjalnym.

Powróćmy do układu złożonego z jednej cząstki, tym razem oddziaływującej z otoczeniem. Funkcję Lagrange'a musimy uzupełnić o wyraz odpowiedzialny za oddziaływanie. Wyraz ten powinien zależeć od współrzędnych i jest dany pewną funkcją zmiennych przestrzennych $U(\vec{r})$:

$$L(\vec{r}, \dot{\vec{r}}, t) = T(\dot{\vec{r}}) - U(\vec{r})$$

Wstawiając ją do równań Lagrange'a uzyskujemy:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{r}_i} - \frac{\partial L}{\partial r_i} = m\ddot{r}_i + \frac{\partial U(\vec{r})}{\partial r_i} = 0$$

lub

$$m\ddot{\vec{r}} = -\nabla U(\vec{r})$$

Rozpoznajemy tu równanie Newtona w polu sił danych potencjałem $U(\vec{r})$. Uzyskaliśmy więc interpretację składnika funkcji Lagrange'a odpowiedzialnego za oddziaływanie z polem zewnętrznym. Jest nim energia potencjalna przyjęta ze znakiem przeciwnym.

Funkcja Lagrange'a układu wielu cząstek.

Uogólnienie uzyskanego wyniku na układ wielu cząstek przeprowadzamy podobnie jak dla układu cząstek nie oddziaływujących, uwzględniając w energii potencjalnej oddziaływanie cząstek z polem zewnętrznym oraz pomiędzy sobą:

$$L = T - U$$

T – stanowi całkowitą energię kinetyczną cząstek tworzących układ, a U – sumę energii potencjalnych cząstek:

$$L = \sum_{a=1}^N T_a - U(\vec{r}_1, \dots, \vec{r}_N) = \sum_{a=1}^N \frac{m_a}{2} \dot{\vec{r}}_a^2 - U(\vec{r}_1, \dots, \vec{r}_N)$$

Transformacja ze współrzędnych kartezjańskich do współrzędnych uogólnionych.

Energię kinetyczną układu cząstek najłatwiej zapisać we współrzędnych kartezjańskich:

$$T = \sum_{a=1}^N \sum_{i=1}^3 \frac{m_a}{2} \dot{x}_{ai}^2.$$

Jednakże potencjał posiada często symetrię preferującą inny układ współrzędnych. Dlatego często musimy dokonać transformacji energii kinetycznej. Dla ułatwienia dyskusji zmienimy oznaczenia zmiennych:

$$x_{ai} = x_{\alpha}, \quad \alpha = 1, \dots, 3N$$

$$T = \sum_{\alpha=1}^{3N} \frac{m_{\alpha}}{2} \dot{x}_{\alpha}^2.$$

Współrzędne układu kartezjańskiego wyrażamy przez współrzędne uogólnione

$$x_{\alpha} = f_{\alpha}(q_1, \dots, q_f)$$

$$\dot{x}_{\alpha} = \sum_{\beta=1}^f \frac{\partial f_{\alpha}}{\partial q_{\beta}} \dot{q}_{\beta}$$

wstawiamy to do wyrażenia na energię kinetyczną

$$T = \sum_{\alpha=1}^{3N} \frac{m_{\alpha}}{2} \left(\sum_{\beta=1}^f \frac{\partial f_{\alpha}}{\partial q_{\beta}} \dot{q}_{\beta} \right)^2 = \sum_{\alpha=1}^{3N} \sum_{\beta=1}^f \sum_{\beta'=1}^f \frac{m_{\alpha}}{2} \frac{\partial f_{\alpha}}{\partial q_{\beta}} \frac{\partial f_{\alpha}}{\partial q_{\beta'}} \dot{q}_{\beta} \dot{q}_{\beta'} = \sum_{\beta, \beta'=1}^f t_{\beta\beta'} \dot{q}_{\beta} \dot{q}_{\beta'}$$

gdzie

$$t_{\beta\beta'} = \sum_{\alpha=1}^{3N} \frac{m_{\alpha}}{2} \frac{\partial f_{\alpha}}{\partial q_{\beta}} \frac{\partial f_{\alpha}}{\partial q_{\beta'}} = t_{\beta\beta'}(\{q_{\beta}\}).$$

Wyrażone powyższymi wzorami formalne przejście do innego układu zmiennych wygląda dość skomplikowanie. Nie zawsze jednak konieczny będzie taki sposób postępowania. Natomiast zwróćmy uwagę na uzyskaną ogólną własność: energia kinetyczna jest formą kwadratową prędkości uogólnionych o współczynniki będących funkcjami współrzędnych uogólnionych.

Funkcja Lagrange'a dla cząstki swobodnej we współrzędnych sferycznych i cylindrycznych.

Dokonajmy dla ilustracji przejścia z układu kartezjańskiego do dwóch najczęściej stosowanych krzywoliniowych układów współrzędnych – cylindrycznego i sferycznego.

1. Współrzędne cylindryczne

$$x = r \cos \varphi$$

$$y = r \sin \varphi$$

$$z = z$$

Energia kinetyczna napisana we współrzędnych kartezjańskich ma postać:

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

Występujące w niej prędkości przeprowadzamy do współrzędnych cylindrycznych

$$\dot{x} = \dot{r} \cos \varphi - \dot{\varphi} r \sin \varphi$$

$$\dot{y} = \dot{r} \sin \varphi + \dot{\varphi} r \cos \varphi$$

$$\begin{aligned} T &= \frac{m}{2} \left(\dot{r}^2 \cos^2 \varphi - 2\dot{r}\dot{\varphi} r \sin \varphi \cos \varphi + \dot{\varphi}^2 r^2 \sin^2 \varphi + \dot{r}^2 \sin^2 \varphi + 2\dot{r}\dot{\varphi} r \sin \varphi \cos \varphi + \dot{\varphi}^2 r^2 \cos^2 \varphi + \dot{z}^2 \right) = \\ &= \frac{m}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2 + \dot{z}^2) \end{aligned}$$

Ponieważ energia potencjalna dla cząstki swobodnej jest równa 0, funkcja Lagrange'a zawiera wyłącznie energię kinetyczną:

$$L(r, \varphi, z, \dot{r}, \dot{\varphi}, \dot{z}, t) = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\varphi}^2 + \dot{z}^2)$$

2. Współrzędne sferyczne

$$x = r \cos \varphi \sin \vartheta$$

$$y = r \sin \varphi \sin \vartheta$$

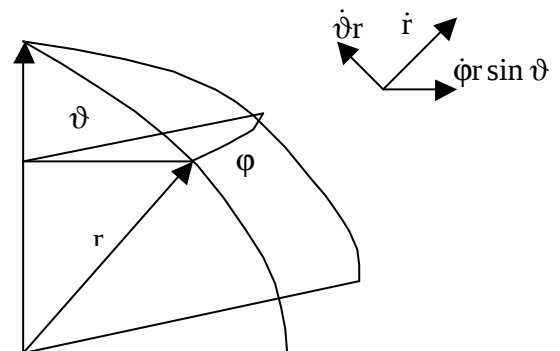
$$z = r \cos \vartheta$$

Możemy podobnie jak poprzednio dokonać podstawienia. Wymaga to trochę dłuższego rachunku niż przy przejściu do współrzędnych cylindrycznych. Posłużymy się zatem inną metodą, dzięki której wyrażenie na kwadrat prędkości można uzyskać bez żadnych rachunków. Wektor prędkości w każdym punkcie przestrzeni rozkładamy na trzy ortogonalne lecz nie kartezjańskie składowe:

$$\dot{\vec{r}} = (\dot{r}, \dot{\vartheta} r, \dot{\varphi} r \sin \vartheta).$$

Otrzymujemy:

$$T = \frac{m}{2} (\dot{r}^2 + \dot{\varphi}^2 r^2 \sin^2 \vartheta + \dot{\vartheta}^2 r^2)$$



Przykładowe zastosowanie formalizmu Lagrange'a.

Cząstka w polu centralnego potencjału.

O wyborze współrzędnych krzywoliniowych często decyduje symetria potencjału, który odejmujemy od energii kinetycznej przy tworzeniu Lagrangianu. W szeregu problemów spotykamy się z potencjałem o symetrii sferycznej zależnym wyłącznie od odległości od wyróżnionego punktu w przestrzeni, tzw. potencjałem centralnym (np. planeta w obecności potencjału grawitacyjnego gwiazdy). Pracujemy wówczas we współrzędnych sferycznych. Lagrangian ma postać:

$$L = \frac{m}{2} (\dot{r}^2 + \dot{\varphi}^2 r^2 \sin^2 \vartheta + \dot{\vartheta}^2 r^2) - V(r),$$

$$\text{Z równań Lagrange'a: } \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} - \frac{\partial L}{\partial q_\alpha} = 0$$

Powinniśmy otrzymać układ równań ruchu. Liczymy odpowiednie pochodne cząstkowe:

$$\frac{\partial L}{\partial \dot{r}} = m\dot{r}, \quad \frac{\partial L}{\partial \dot{\varphi}} = m\dot{\varphi} r^2 \sin^2 \vartheta, \quad \frac{\partial L}{\partial \dot{\vartheta}} = mr^2 \dot{\vartheta}$$

$$\frac{\partial L}{\partial r} = m(\dot{\varphi}^2 r \sin^2 \vartheta + \dot{\vartheta}^2 r) - \frac{\partial V}{\partial r}, \quad \frac{\partial L}{\partial \varphi} = 0, \quad \frac{\partial L}{\partial \vartheta} = m\dot{\varphi}^2 r^2 \cos \vartheta \sin \vartheta$$

Po wstawieniu ich do równań Lagrange'a otrzymujemy układ równań:

$$m\ddot{r} - m(\dot{\varphi}^2 r \sin^2 \vartheta + \dot{\vartheta}^2 r) + \frac{\partial V}{\partial r} = 0$$

$$mr^2 \sin^2 \vartheta \ddot{\varphi} + 2mr \sin^2 \vartheta \dot{r} \dot{\varphi} + 2mr^2 \cos \vartheta \sin \vartheta \dot{\vartheta} \dot{\varphi} = 0$$

$$mr^2 \ddot{\vartheta} + 2mr \dot{\vartheta} \dot{r} - mr^2 \dot{\varphi}^2 \cos \vartheta \sin \vartheta = 0$$

Uzyskaliśmy układ trzech równań różniczkowych drugiego stopnia względem czasu na szukaną trajektorię $(r(t), \varphi(t), \vartheta(t))$.

Analityczne rozwiązanie tego układu równań jest niestety skomplikowane. Znacznie łatwiej wykonać to rozwiązując problem numerycznie.

Wahadło matematyczne.

Żeby naprawdę docenić zalety stosowania formalizmu Lagrange'a rozważmy prosty przykład, w którym występujące więzy utrudniają wydobycie równań ruchu z zasad dynamiki Newtona.

Rozważmy układ jaki stanowi płaskie wahadło matematyczne, czyli punkt materialny o masie M zawieszony na nierozciągliwej nici o długości R . Układ taki ma tylko jeden stopień swobody, który we współrzędnych biegunowych stanowi kąt odchylenia wahadła od pionu.

W tych współrzędnych energia kinetyczna układu ma postać:

$$T = \frac{M}{2} R^2 \dot{\varphi}^2$$

Energia potencjalna liczona względem punktu zawieszenia wahadła dana jest wyrażeniem:

$$U = -MgR \cos \varphi$$

Zatem funkcja Lagrange'a:

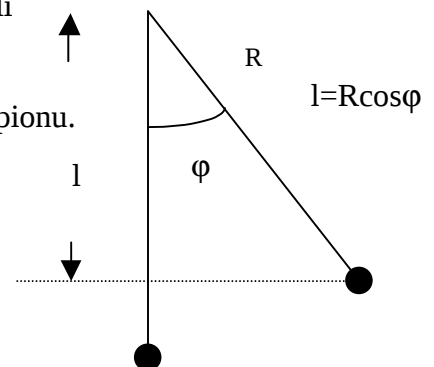
$$L = T - U = \frac{M}{2} R^2 \dot{\varphi}^2 + MgR \cos \varphi$$

Równanie Lagrange'a generuje równanie ruchu:

$$MR^2 \ddot{\varphi} + MgR \sin \varphi = 0,$$

które po uproszczeniu przyjmuje postać:

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

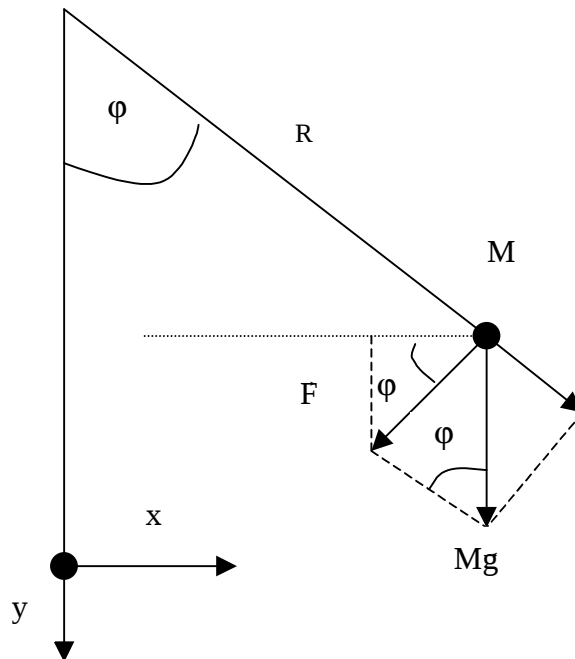


Oczywiście to samo równanie ruchu powinniśmy uzyskać bezpośrednio z równania Newtona. Przyjrzyjmy się takiemu rozumowaniu.

Przy założeniu płaskiego ruchu położenie wahadła opisywane jest dwuwymiarowym wektorem wodzącym $\vec{r}(t)$. Równanie ruchu

$$M\ddot{\vec{r}} = \vec{F} = M\vec{g} + \vec{F}_r$$

zawiera siłę ciężkości i siłę reakcji więzów. Wartość tej ostatniej siły nie jest znana, ale z warunku przenoszenia tej siły przez elastyczną nić znamy kierunek jej działania i wiemy, że będzie równoważył równoległą do nici składową siły ciężkości.



W rezultacie wypadkowa siła działająca na ciało jest prostopadła do nici a jej wartość:

$$F = Mg \sin \varphi$$

Dwuwymiarowe równanie Newtona rozpisujemy na składowe kartezjańskie:

$$M\ddot{x} = -Mg \sin \varphi \cos \varphi$$

$$M\ddot{y} = Mg \sin^2 \varphi$$

Przechodzimy teraz do współrzędnych biegunowych

$$x = R \sin \varphi$$

$$y = R \cos \varphi$$

$$\dot{x} = R \dot{\varphi} \cos \varphi$$

$$\ddot{x} = R \ddot{\varphi} \cos \varphi - R \dot{\varphi}^2 \sin \varphi$$

$$\dot{y} = -R \dot{\varphi} \sin \varphi$$

$$\ddot{y} = -R \ddot{\varphi} \sin \varphi - R \dot{\varphi}^2 \cos \varphi$$

Uzyskane pochodne wstawiamy do obu równań ruchu

$$M(R \ddot{\varphi} \cos \varphi - R \dot{\varphi}^2 \sin \varphi) = -Mg \sin \varphi \cos \varphi$$

$$M(R \ddot{\varphi} \sin \varphi + R \dot{\varphi}^2 \cos \varphi) = -Mg \sin^2 \varphi$$

Górne równanie mnożymy przez $\cos \varphi$ dolne przez $\sin \varphi$ i dodajemy stronami. Wynik po redukcji niektórych elementów przyjmuje postać:

$$R \ddot{\varphi} = -g \sin \varphi$$

Uzyskaliśmy to samo równanie ruchu, ale porównanie długości rachunków wypada na korzyść formalizmu Lagrange'a.

Podsumowanie II.

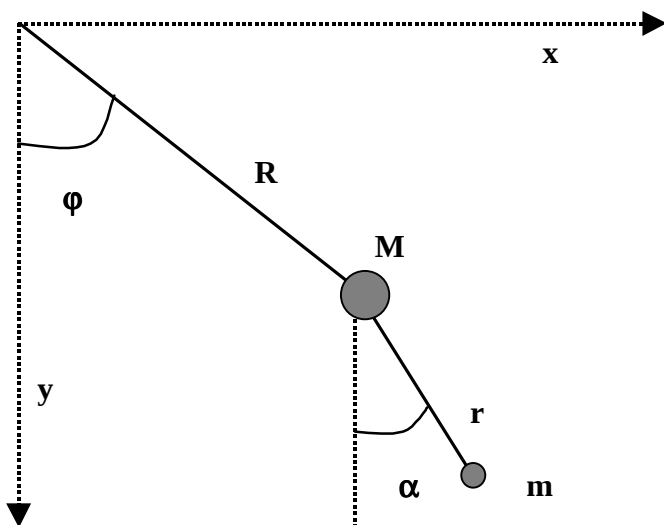
- Wprowadziliśmy pojęcie układu inercjalnego, założyliśmy równoważność praw mechaniki we wszystkich układach inercjalnych.
- Wprowadziliśmy pojęcie transformacji Galileusza opisującej przejścia pomiędzy układami inercjalnymi w mechanice klasycznej.
- Korzystając z jednorodności i izotropowości przestrzeni oraz jednorodności czasu uzyskaliśmy wyrażenie na funkcję Lagrange'a dla cząstki swobodnej. Sprawdziliśmy, że równania Lagrange'a są równoważne równaniom Newtona.
- Uogólniliśmy funkcję Lagrange'a na przypadek cząstki w polu zewnętrznego potencjału oraz układu wielu cząstek oddziaływujących z polem zewnętrznym i pomiędzy sobą.
- Pokazaliśmy, że przy przejściu do dowolnych współrzędnych uogólnionych energia kinetyczna cząstki jest biliniową formą prędkości uogólnionych.
- Napisaliśmy funkcje Lagrange'a we współrzędnych cylindrycznych i sferycznych. Znaleźliśmy równania ruchu we współrzędnych sferycznych dla problemu cząstki w polu sił centralnych.
- Na przykładzie wahadła matematycznego porównaliśmy sposób pozyskiwania równań ruchu z formalizmu Lagrange'a i z zasad dynamiki Newtona.

III. Bardziej zaawansowane zastosowania formalizmu Lagrange'a.

Na poprzednim wykładzie poznaliśmy najprostsze zastosowania formalizmu Lagrange'a. Mam nadzieję, że już one pozwoliły pokazać jego zalety. Oba problemy można było rozwiązać łatwiej i nieco szybciej niż startując z równań Newtona. Dla osób, które jeszcze nie są przekonane do tego formalizmu przygotowałem kilka przykładów wykorzystania formalizmu Lagrange'a.

Wahadło złożone.

Powróćmy do wahadła matematycznego o masie M i długości R ale do niego przyczepmy drugie wahadło o masie m i długości r . Problem się znacznie komplikuje. Znalazienie sił reakcji więzów staje się znaczącym problemem rachunkowym. W takich przypadkach formalizm Lagrange'a stanowi znakomite narzędzie pracy usuwające kłopoty.



Dodatkowe ciało wnosi dodatkowy stopień swobody. Opiszemy go przy pomocy kąta α określającego odchylenie nici r od pionu. Położenie masy m możemy zapisać we współrzędnych kartezjańskich jako:

$$x = R \sin \varphi + r \sin \alpha$$

$$y = R \cos \varphi + r \cos \alpha$$

Energia kinetyczna ciała o masie m :

$$T_m = \frac{m}{2} (R^2 \dot{\varphi}^2 + r^2 \dot{\alpha}^2 + 2Rr \cos(\varphi - \alpha) \dot{\varphi} \dot{\alpha})$$

Energia potencjalna

$$U_m = -mg(R \cos \varphi + r \cos \alpha)$$

Funkcja Lagrange'a dla drugiego ciała

$$L_m = \frac{m}{2} (R^2 \dot{\varphi}^2 + r^2 \dot{\alpha}^2 + 2Rr \cos(\varphi - \alpha) \dot{\varphi} \dot{\alpha}) + mg(R \cos \varphi + r \cos \alpha)$$

Po dodaniu funkcji Lagrange'a dla obu ciał uzyskujemy funkcję Lagrange'a całego układu:

$$L = L_M + L_m = T_M - U_M + T_m - U_m = \frac{M}{2} R^2 \dot{\varphi}^2 + MgR \cos \varphi +$$

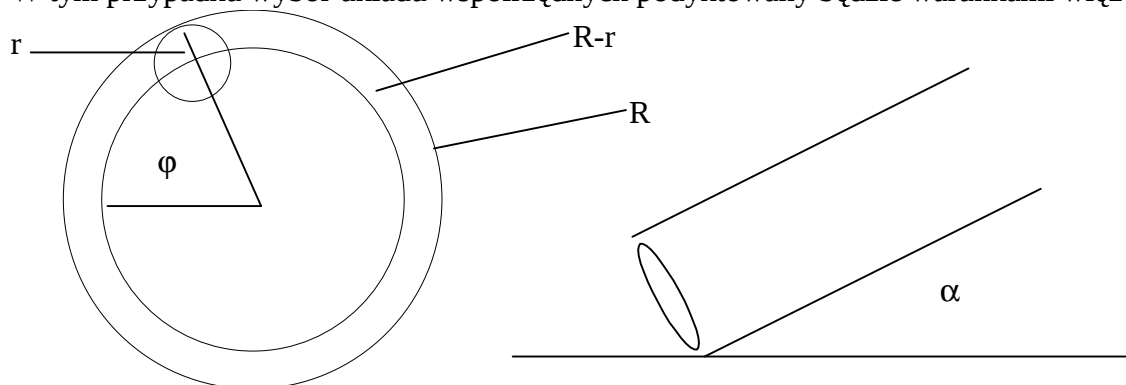
$$+ \frac{m}{2} (R^2 \dot{\varphi}^2 + r^2 \dot{\alpha}^2 + 2Rr \cos(\varphi - \alpha) \dot{\varphi} \dot{\alpha}) + mg(R \cos \varphi + r \cos \alpha)$$

Wzór ten można uporządkować i wyliczyć układ dwóch równań ruchu odpowiadających obu stopniom swobody.

Przykład zastosowania formalizmu Lagrange'a do opisu ruchu bryły sztywnej.

Ponieważ zagadnienia mechaniki klasycznej nie ograniczają się do układów punktów materialnych, rozwiążmy zadanie, w którym dyskutować będziemy problem niepunktowego ciała.

Poszukajmy równań ruchu dla jednorodnej kulki o masie m i promieniu r , toczącej się bez poślizgu po wewnętrznej stronie poboczniczy walca o promieniu R , umieszczonego w polu sił ciężkości tak, że jego oś nachylona jest pod kątem α do poziomu. Dla utrudnienia założmy, że promień kulki r nie jest zanedbywalnie mały w porównaniu z promieniem cylindra R . W tym przypadku wybór układu współrzędnych podyktowany będzie warunkami więzów.



Układ ma dwa stopnie swobody. We współrzędnych cylindrycznych są nimi

$$U_1 = -m_2 g(d + r \cos \varphi).$$

W energii kinetycznej musimy uwzględnić energię ruchu postępowego i obrotowego kulki

$$T = \frac{I\omega^2}{2} + \frac{mv^2}{2}$$

gdzie I jest momentem bezwładności kulki:

$$I = \frac{2}{5}mr^2$$

$$L = T_A(q_A, \dot{q}_A, t) - U(q_A, q_B(t))$$

Prędkość ruchu postępowego rozdzielamy na składową wzdłuż osi walca i prostopadłą do niej:

$$V_z = \dot{z} \quad \text{i} \quad V_\varphi = (R - r)\dot{\varphi}.$$

Z warunku toczenia bez poślizgu wyznaczamy związek pomiędzy prędkością kątową ruchu obrotowego kulki i prędkością ruchu postępowego. Prędkość ruchu obrotowego musimy również rozłożyć na dwie składowe:

$$\omega_z r = \dot{z}, \quad \omega_\varphi r = -R\dot{\varphi}$$

Po podstawieniu uzyskanych wyrażeń na składowe prędkości do wzoru na energię kinetyczną uzyskujemy:

$$T = \frac{m(\dot{z}^2 + R^2\dot{\varphi}^2)}{5} + \frac{m(\dot{z}^2 + (R - r)^2\dot{\varphi}^2)}{2} = \frac{7m}{10}\dot{z}^2 + m\left(\frac{R^2}{5} + \frac{(R - r)^2}{2}\right)\dot{\varphi}^2$$

Energia potencjalna pola sił ciężkości związana jest z wysokością ponad założony poziom odniesienia, który umieścimy tak, by wyrażenie na energię było najprostsze. Jako punkt odniesienia przyjmujemy najniższy punkt osi walca.

$$V(z, \varphi) = mg(z \sin \alpha + (R - r) \sin \varphi \cos \alpha)$$

Łącząc energię kinetyczną i potencjalną uzyskujemy funkcję Lagrange'a:

$$L = T - V = \frac{7m}{10} \dot{z}^2 + m \left(\frac{R^2}{5} + \frac{(R-r)^2}{2} \right) \dot{\phi}^2 - mg(z \sin \alpha + (R-r) \sin \phi \cos \alpha)$$

$$\frac{\partial L}{\partial \dot{z}} = \frac{7m}{5} \dot{z}, \frac{\partial L}{\partial \dot{\phi}} = m \left(\frac{2R^2}{5} + (R-r)^2 \right) \dot{\phi}$$

$$\frac{\partial L}{\partial z} = -mg \sin \alpha, \frac{\partial L}{\partial \phi} = -mg(R-r) \cos \phi \cos \alpha$$

I równania ruchu:

$$\frac{7m}{5} \ddot{z} + mg \sin \alpha = 0$$

$$m \left(\frac{2R^2}{5} + (R-r)^2 \right) \ddot{\phi} + mg(R-r) \cos \phi \cos \alpha = 0$$

W rozważanym przykładzie równania na współrzędne odpowiadające obu stopniom swobody są niezależne. Mówimy o separacji zmiennych. W kierunku równoległym do osi walca mamy toczenie kulki po równi pochyłej a w kierunku prostopadłym, toczenie kulki po wnętrzu okręgu. Tej separacji mogliśmy oczekiwać i oba ruchy potraktować niezależnie.

Możliwość separacji ruchu pojawia się zawsze kiedy zarówno energia kinetyczna jak i potencjalna dadzą się przedstawić w postaci sumy składników, z których każdy zawiera tylko jedną zmienną.

Wydzielenie podukładu z układu cząstek oddziaływujących.

Rozważmy teraz następujący przykład. Załóżmy, że układ udaje się podzielić na **dwa** podukłady A i B. Wówczas funkcja Lagrange'a całego układu zawiera sumę Lagrangianów obu podukładów, oraz potencjał oddziaływania pomiędzy nimi, zależny wyłącznie od współrzędnych uogólnionych obu podukładów:

$$L = L_A(q_A, \dot{q}_A, t) + L_B(q_B, \dot{q}_B, t) - U(q_A, q_B)$$

Jeżeli ustalimy rzeczywisty ruch podukładu B, zadając odpowiednie trajektorie:

$$q_B = q_B(t)$$

Współrzędne uogólnione podukładu B przestają być swobodnymi parametrami układu.

Lagrangian układu B możemy potraktować jako funkcję czasu i potraktować ją jako

pochoďna zupełna po czasie jej całki nieoznaczonej

$$L_B = L_B(t) = \frac{d}{dt} \int L_B(t) dt$$

Można ją wtedy usunąć z Lagrangianu.

Również energia potencjalna staje się wówczas funkcją współrzędnych uogólnionych podukładu A i czasu. Tak więc otrzymujemy:

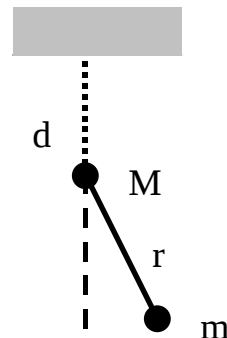
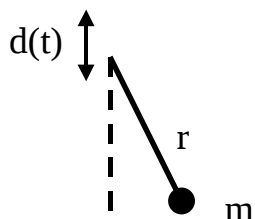
$$L = L_A(q_A, \dot{q}_A, t) - U(q_A, q_B(t))$$

funkcję Lagrange'a układu A w polu zewnętrznego potencjału wywołanego oddziaływaniem z podukładem B.

Zilustrujmy to przykładem.

Chcemy rozwiązać problem wahadła matematycznego, którego punkt zawieszenia wykonuje pionowe drgania zadane funkcją: $d(t) = d_0 + C \cos \omega t$.

Startując z równań Newtona z całą pewnością będziemy mieli wątpliwości jakie siły działają na wahadło. Użyjemy formalizmu Lagrange'a i przedstawionego powyżej tricku z podziałem układu na części.



Rozważamy układ złożony z punktowej masy m_1 zawieszony na sprężynie o długości d_0 i współczynniku sprężystości k . Masa m_1 może wykonywać pionowe ruchy. Do masy m_1 podwieszone jest

wahadło matematyczne o długości r i masie m_2 mogące poruszać się w wyróżnionej płaszczyźnie pionowej. Znajdźmy Lagrangian tego układu.

Układ posiada 2 stopnie swobody: długość sprężyny d i kąt odchylenia wahadła od pionu φ .

Położenie pierwszego ciała wyznacza długość sprężyny d , położenie drugiego rozkładamy na dwie składowe pionową y i poziomą x :

$$y = d + r \cos \varphi, \quad \dot{y} = \dot{d} - r \sin \varphi \dot{\varphi}$$

$$x = r \sin \varphi, \quad \dot{x} = r \cos \varphi \dot{\varphi}$$

Wyrażamy energie kinetyczne obu ciał przez prędkości uogólnione odpowiadające stopniom swobody:

$$T_1 = \frac{M}{2} \dot{d}^2$$

$$T_2 = \frac{m}{2} \left((\dot{d} - r \sin \varphi \dot{\varphi})^2 + r^2 \cos^2 \varphi \dot{\varphi}^2 \right) = \frac{m}{2} (\dot{d}^2 - 2\dot{d}\dot{\varphi} r \sin \varphi + r^2 \dot{\varphi}^2)$$

Podobnie energie potencjalne:

$$U_1 = \frac{k}{2} (d - d_0)^2 - Mgd$$

$$U_2 = -Mg(d + r \cos \varphi)$$

Otrzymujemy:

$$L = T_1 + T_2 - U_1 - U_2 =$$

$$= \frac{M}{2} \dot{d}^2 + \frac{m}{2} (\dot{d}^2 - 2\dot{d}\dot{\varphi} r \sin \varphi + r^2 \dot{\varphi}^2) - \frac{k}{2} (d - d_0)^2 + Mgd + mg(d + r \cos \varphi)$$

Rozważany układ możemy podzielić na dwie części. Niech część A stanowi wahadło, a część B masa M na sprężynie. Ustalmy ruch części B układu, zakładając, że masa m wykonuje ustalone drgania dane funkcją:

$$d(t) = d_0 + C \cos \omega t$$

Przy tym założeniu możemy odrzucić Lagrangian pierwszej cząstki oraz w pozostałej części zastąpić zmienną niezależną d funkcją czasu $d(t)$.

$$L = + \frac{m}{2} (\dot{d}(t)^2 - 2\dot{d}(t)\dot{\varphi} r \sin \varphi + r^2 \dot{\varphi}^2) + mg(d(t) + r \cos \varphi) =$$

$$= + \frac{m}{2} (C^2 \omega^2 \sin^2 \omega t + 2C\omega \dot{\varphi} \sin \omega t r \sin \varphi + r^2 \dot{\varphi}^2) + mg(d_0 + C \cos \omega t + r \cos \varphi)$$

Możemy dokonać dalszych zabiegów na Lagrangianie polegających na usunięciu wszystkich składników dających się przedstawić w postaci pochodnej zupełnej po czasie. Pozostaje jego równoważna postać:

$$L' = + \frac{m}{2} (2C\omega\dot{\phi}\sin\omega r\sin\phi + r^2\dot{\phi}^2) + mgr\cos\phi$$

Układ posiada oczywiście tylko jeden stopień swobody odpowiadający współrzędnej ϕ .

Równania ruchu w nieinercjalnych układach odniesienia.

Zasadę najmniejszego działania i wynikające z niej równania Lagrange'a wprowadziliśmy przed zdefiniowaniem pojęcia układu inercjalnego. Formalizm Lagrange'a jest więc słuszny w dowolnych, również nieinercjalnych układach odniesienia. Wystarczy znać funkcję Lagrange'a a z niej można otrzymać równania ruchu. Funkcję Lagrange'a na razie potrafimy skonstruować tylko w układzie inercjalnym. W układzie nieinercjalnym uzyskamy ją przechodząc z układu inercjalnego. Załóżmy, że w układzie inercjalnym który oznaczymy przez Σ znajduje się cząstka oddziaływująca z zewnętrznym potencjałem $U(\vec{r})$. Funkcja Lagrange'a w tym układzie odniesienia ma postać:

$$L = \frac{m}{2} \vec{v}^2 - U(\vec{r})$$

Oznaczmy przez Σ' drugi układ odniesienia poruszający się względem układu Σ z prędkością $\vec{V}(t)$. Układ ten będzie również inercjalny, jeżeli prędkość $\vec{V}(t)$ będzie stała w czasie. Założenie zależności czasowej prędkości względnej układów prowadzi do układu nieinercjalnego.

Jeżeli w układzie Σ' cząstka porusza się z prędkością $\vec{v}'$, to w układzie Σ posiada prędkość $\vec{v} = \vec{v}' + \vec{V}(t)$

Wstawmy to wyrażenie do funkcji Lagrange'a :

$$L = \frac{m}{2} (\vec{v}' + \vec{V}(t))^2 - U = \frac{m}{2} \vec{v}'^2 + m\vec{v}'\vec{V}(t) + \frac{m}{2} \vec{V}(t)^2 - U$$

Zauważmy, że trzeci składnik sumy stanowi pochodną zupełną pewnej funkcji i może zostać z funkcji Lagrange'a usunięty.

Składnik drugi przekształcimy korzystając ze związku:

$$m\vec{v}'\vec{V}(t) = \frac{d}{dt} (m\vec{r}'\vec{V}(t)) - m\vec{r}' \frac{d}{dt} \vec{V}(t)$$

Pierwszą część tego wyrażenia można również usunąć z funkcji Lagrange'a. Pozostaje:

$$L = \frac{m}{2} \vec{v}'^2 - m\vec{r}'\dot{\vec{V}}(t) - U$$

Możemy teraz uzyskać równania ruchu:

$$m\ddot{\vec{r}}' + m\dot{\vec{V}}(t) + \nabla U = 0$$

lub

$$m\ddot{\vec{r}}' = -m\dot{\vec{V}}(t) - \nabla U$$

Zauważmy, że w równaniu ruchu pojawiła się proporcjonalna do masy cząstki jednorodna siła skierowana przeciwnie do kierunku przyspieszenia układu nieinercjalnego. Możemy ją nazwać siłą bezwładności. Wynik wydaje się dość oczywisty. Przejdźmy zatem do analizy mniej oczywistego problemu. Rozważmy teraz układ Σ , którego początek pokrywa się z początkiem układu Σ' , lecz obracający się względem niego z prędkością kątową $\vec{\Omega}$ (zależną od czasu). Wybierzmy osie z i z' równoległe do prędkości kątowej $\vec{\Omega}$ i znajdziemy związki pomiędzy prędkościami w obu układach.

Składowe wektora wodzącego $\vec{r}$ w płaszczyźnie x, y wyrażamy przez odległość r od początków układów i kąt φ w układzie Σ :

$$x = r \sin \varphi$$

$$y = r \cos \varphi$$

Jeżeli w czasie dt cząstki przebywają w układzie Σ odcinek (dx,dy), to w układzie Σ' przebędą dodatkowe odcinki wynikające ze zwiększenia kąta φ o Ωdt .

$$dx' = dx + r \sin(\varphi + \Omega dt) - r \sin \varphi = dx + r \Omega dt \cos \varphi = dx + y \Omega dt$$

$$dy' = dy + r \cos(\varphi + \Omega dt) - r \cos \varphi = dy - r \Omega dt \sin \varphi = dy - x \Omega dt$$

Po podzieleniu przez dt otrzymujemy:

$$\dot{\vec{r}}' = \dot{\vec{r}} + \Omega(y, -x, 0)$$

Ponieważ układ współrzędnych wybraliśmy tak, żeby

$$\vec{\Omega} = (0, 0, \Omega)$$

zatem

$$\vec{\Omega} \times \vec{r} = \begin{vmatrix} 0 & 0 & \Omega \\ x & y & z \\ \vec{i} & \vec{j} & \vec{k} \end{vmatrix} = \Omega x \vec{j} - \Omega y \vec{i} = \Omega(-y, x, 0).$$

Poprzednie wyrażenie można zapisać jako:

$$\dot{\vec{r}}' = \dot{\vec{r}} - \vec{\Omega} \times \vec{r}$$

Wstawmy ten wynik do odpowiedniego wyrażenia na funkcję Lagrange'a. Otrzymujemy:

$$L = \frac{m}{2} \vec{v}'^2 + m \vec{v}' \vec{V}(t) + \frac{m}{2} \vec{V}(t)^2 - U = \frac{m}{2} (\dot{\vec{r}} - \vec{\Omega} \times \vec{r})^2 + m (\dot{\vec{r}} - \vec{\Omega} \times \vec{r}) \vec{V}(t) - U \text{ zatem}$$

$$L = \frac{m}{2} \dot{\vec{r}}^2 - m \dot{\vec{r}} \cdot \vec{\Omega} \times \vec{r} + \frac{m}{2} (\vec{\Omega} \times \vec{r})^2 - m \vec{r} \dot{\vec{V}}(t) - m \vec{r} (\vec{V}(t) \times \vec{\Omega}) - U$$

Jest to funkcja Lagrange'a dla cząstki w układzie nieinercyjnym, poruszającym się względem pewnego układu inercyjnego z prędkością liniową $\vec{V}(t)$ i obracającym się względem niego z prędkością kątową $\vec{\Omega}(t)$.

Napiszmy równania ruchu dla tej cząstki.

Dla uproszczenia zapisu wprowadzę pojęcie pochodnej kierunkowej, którą oznaczę jako

$$\frac{\partial}{\partial \vec{r}}.$$

Będzie oznaczać wektor o składowych

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right).$$

Dla pochodnej kierunkowej po wektorze wodzącym jest to po prostu operator ∇ , a dla pochodnej po prędkości odpowiednik $\nabla_{\dot{\vec{r}}}$.

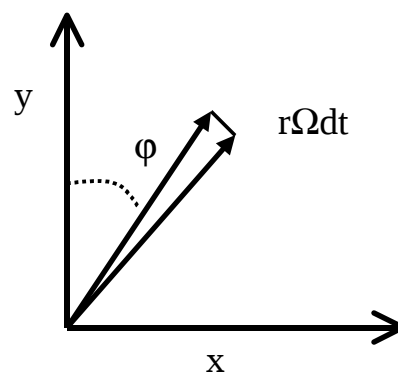
W tym zapisie równania Lagrange'a mają postać:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\vec{r}}} - \frac{\partial L}{\partial \vec{r}} = 0.$$

Wykorzystajmy je do uzyskanej funkcji Lagrange'a.

$$\frac{\partial L}{\partial \dot{\vec{r}}} = m \dot{\vec{r}} - m \vec{\Omega} \times \vec{r}$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\vec{r}}} = m \ddot{\vec{r}} - m \dot{\vec{\Omega}} \times \vec{r} - m \vec{\Omega} \times \dot{\vec{r}}$$



$$\frac{\partial L}{\partial \vec{r}} = -m\dot{\vec{r}} \times \vec{\Omega} - m(\vec{\Omega} \times \vec{r}) \times \vec{\Omega} - m\dot{\vec{V}}(t) - m(\vec{V}(t) \times \vec{\Omega}) - \nabla U$$

Równania ruchu przyjmują postać:

$m\ddot{\vec{r}} = m\dot{\vec{\Omega}} \times \vec{r} + m\vec{\Omega} \times \dot{\vec{r}} - m\dot{\vec{r}} \times \vec{\Omega} - m(\vec{\Omega} \times \vec{r}) \times \vec{\Omega} - m\dot{\vec{V}}(t) - m(\vec{V}(t) \times \vec{\Omega}) - \nabla U$ co po uporządkowaniu daje:

$$m\ddot{\vec{r}} = -\nabla U - m\dot{\vec{V}}(t) + m\dot{\vec{\Omega}} \times \vec{r} + 2m\vec{\Omega} \times \dot{\vec{r}} - m(\vec{\Omega} \times \vec{r}) \times \vec{\Omega} - m(\vec{V}(t) \times \vec{\Omega})$$

Tak więc oprócz siły wynikającej z potencjału U pojawia się podobnie jak poprzednio jednorodna siła bezwładności wynikająca z przyspieszonego ruchu prostoliniowego układu.

Dodatkowe wyrażenia mają również charakter sił bezwładności wynikających z ruchu obrotowego układu. Pierwsza z nich wynika ze zmian prędkości kątowych, następna ma charakter siły Coriolisa, a druga siły odśrodkowej. Zwróćmy uwagę na kierunki i zwroty kolejnych składników sił bezwładności.

Łatwo sprawdzić, że w przypadku ciała spoczywającego w poruszającym się ruchem obrotowym układzie odniesienia Kierunek tej ostatniej siły pokrywa się z promieniem wodzącym cząstki w układzie. Kierunek siły Coriolisa jest prostopadły do osi obrotu układu nieinercyjnego oraz wektora prędkości.

Podsumowanie III.

- Zastosowaliśmy formalizm Lagrange'a do podwójnego wahadła matematycznego.
- Następnie do uzyskania równań ruchu bryły sztywnej.
- Nauczyliśmy się jak sobie radzić z trajektorią narzuconą na jedną ze zmiennych przez wydzielenia z układu jego części, zadania jego ruchu i potraktowania jej jako układu zewnętrznego.
- Na zakończenie uzyskaliśmy równanie ruchu w nieinercyjnym układzie odniesienia. Założyliśmy ruch nieinercyjnego układu odniesienia z zależną od czasu prędkością liniową oraz jego obrót ze zmienną prędkością kątową.

IV.

Prawa zachowania, całki ruchu.

Spośród wszystkich wielkości fizycznych, które można obserwować w układzie szczególne znaczenie mają te wielkości, które są stałe w czasie. Mówimy o ich zachowaniu. W różnych układach, czyli w zależności od charakteru potencjałów (sił) zewnętrznych mogą być zachowane różne wielkości. Wiedząc jakie prawa zachowania słuszne są w rozpatrywanym układzie możemy wykorzystać je w celu ułatwienia rozwiązania problemu. Wielkości zachowane w czasie służą nam również do klasyfikacji (określenia) stanu układu.

Zacznijmy od zupełnie ogólnych rozważań, które pozwolą nam określić liczbę niezależnych wielkości fizycznych, dla których można sformułować prawa zachowania. Wielkości te nazywamy całkami ruchu.

Układ o f stopniach swobody jest podczas ruchu opisywany przez $2f$ wielkości będących funkcjami czasu - współrzędnych i prędkości uogólnionych. Szukamy nie zmieniających się w czasie funkcji tych wielkości

$$f(q_\alpha, \dot{q}_\alpha) = \text{const}(t).$$

Zauważmy, że trajektorie, czyli zależności czasowe współrzędnych uogólnionych wyliczamy rozwiązując układ f równań ruchu. Ponieważ równania ruchu są równaniami różniczkowymi drugiego rzędu, do rozwiązań szczególnych wprowadzamy $2f$ warunków początkowych.

Rozwiązania muszą zatem zawierać $2f$ stałych:

$$q_\alpha = q_\alpha(t, c_1, \dots, c_{2f})$$

$$\dot{q}_\alpha = \dot{q}_\alpha(t, c_1, \dots, c_{2f})$$

Traktując powyższe związki jako układ równań możemy go rozwikłać względem stałych c_i i uzyskać $2f$ wielkości, które niewątpliwie od czasu nie zależą, a stanowią funkcje współrzędnych i prędkości uogólnionych.

Jedną z tych stałych usuwamy ustalając chwilę początkową, np.: $c_{2f} = t_0$. Pozostaje $2f-1$ wielkości. Możemy więc sformułować ogólną regułę:

Dla układu zawierającego f stopni swobody można określić $2f-1$ całek ruchu.

Całki ruchu jako wielkości stałe w czasie można wykorzystać przy rozwiązywaniu równań ruchu, można również posługiwać się nimi do określania stanu układu.

Niektóre z nich mają bardziej uniwersalny charakter. Szczególne znaczenie mają całki ruchu bezpośrednio związane z własnościami czasu i przestrzeni – jednorodnością i izotropowością. Mają one szczególną własność - są addytywne. Ich wartość dla układu, który można podzielić na dwa nie oddziałujące podukłady jest sumą ich odpowiednich wartości dla podukładów.

Energia.

Zacznijmy od całki ruchu związanej z jednorodnością czasu. Jeżeli rozważany układ jest izolowany, z warunku jednorodności czasu wynika, że funkcja Lagrange'a nie zawiera jawnej zależności czasowej:

$$\frac{\partial L}{\partial t} = 0.$$

Zatem jej zupełna pochodna po czasie wynika wyłącznie z zależności czasowej współrzędnych i prędkości uogólnionych:

$$\frac{dL}{dt} = \sum_{\alpha} \left(\frac{\partial L}{\partial q_{\alpha}} \dot{q}_{\alpha} + \frac{\partial L}{\partial \dot{q}_{\alpha}} \ddot{q}_{\alpha} \right).$$

Korzystając z równań Lagrange'a przekształcamy ją do:

$$\frac{dL}{dt} = \sum_{\alpha} \left(\dot{q}_{\alpha} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_{\alpha}} + \frac{\partial L}{\partial \dot{q}_{\alpha}} \ddot{q}_{\alpha} \right) = \sum_{\alpha} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_{\alpha}} \dot{q}_{\alpha} \right) = \frac{d}{dt} \sum_{\alpha} \left(\frac{\partial L}{\partial \dot{q}_{\alpha}} \dot{q}_{\alpha} \right)$$

Odejmując od siebie obie strony równania uzyskujemy:

$$\frac{d}{dt} \left(\sum_{\alpha} \frac{\partial L}{\partial \dot{q}_{\alpha}} \dot{q}_{\alpha} - L \right) = 0,$$

skąd wynika:

$$\sum_{\alpha} \frac{\partial L}{\partial \dot{q}_{\alpha}} \dot{q}_{\alpha} - L = E = \text{const}(t)$$

Wielkość E ma wymiar energii i jest całką ruchu dla każdego układu izolowanego. Ponieważ jej niezależność czasową uzyskaliśmy przy założeniu $\frac{\partial L}{\partial t} = 0$, wynik ten możemy

natychmiast uogólnić na układy znajdujące się w polu sił zewnętrznych niezależnych od czasu. Takie układy nazywamy zachowawczymi.

Ponieważ uprzednio pokazaliśmy, że energia kinetyczna układu cząstek jest formą kwadratową prędkości uogólnionych możemy zauważyć, że:

$$\sum_{\alpha} \dot{q}_{\alpha} \frac{\partial L}{\partial \dot{q}_{\alpha}} = \sum_{\alpha} \dot{q}_{\alpha} \frac{\partial T}{\partial \dot{q}_{\alpha}} = 2T,$$

czyli:

$$E = 2T(q, \dot{q}) - L.$$

Ponieważ funkcja Lagrange'a jest różnicą energii kinetycznej i potencjalnej

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

otrzymujemy:

$$E = T(q, \dot{q}) + U(q).$$

Całka ruchu E jest sumą energii kinetycznej i potencjalnej, stanowi więc energię całkowitą układu.

Pęd.

Wydobądźmy teraz całki ruchu związane z jednorodnością przestrzeni. Rozpatrzmy izolowany układ kilku cząstek. Załóżmy, że nie występują w nim potencjały zewnętrzne. W takim przypadku przestrzeń możemy traktować jako jednorodną, co oznacza, że żaden punkt w przestrzeni nie jest wyróżniony.

Rozpatrujemy nieskończenie małe przesunięcie równoległe całego układu i żądamy niezmienniczości funkcji Lagrange'a względem takiego przesunięcia. Jest ona wynikiem jednorodności przestrzeni.

$$\vec{r}_a \rightarrow \vec{r}_a + \vec{\epsilon} \quad , \quad r_{ai} \rightarrow r_{ai} + \epsilon_i$$

$$\delta L = \sum_{ai} \frac{\partial L}{\partial r_{ai}} \epsilon_i = \sum_i \epsilon_i \sum_a \frac{\partial L}{\partial r_{ai}} = 0$$

Ponieważ ϵ_i są dowolne, otrzymujemy warunek:

$$\sum_a \frac{\partial L}{\partial r_{ai}} = 0$$

Korzystając z równań Lagrange'a otrzymujemy:

$$0 = \sum_a \frac{\partial L}{\partial r_{ai}} = \sum_a \frac{d}{dt} \frac{\partial L}{\partial \dot{r}_{ai}} = \frac{d}{dt} \sum_a \frac{\partial L}{\partial \dot{r}_{ai}}$$

Wyrażenie, którego pochodna się zeruje stanowi następną całkę ruchu:

$$P_i = \sum_a \frac{\partial L}{\partial \dot{r}_{ai}} = \text{const}(t).$$

Ponieważ we współrzędnych kartezjańskich

$$L = \sum_{ai} \frac{m_a}{2} \dot{r}_i^2 - U(r_i),$$

dla nowej całki ruchu otrzymujemy:

$$P_i = \sum_a m_a \dot{r}_{ai} \quad \text{lub} \quad \vec{P} = \sum_a m_a \dot{\vec{r}}_a$$

Uzyskana całka ruchu stanowi pęd całkowity układu.

Z drugiej strony ten sam warunek

$$0 = \sum_a \frac{\partial L}{\partial r_{ai}} = - \sum_a \frac{\partial U}{\partial r_{ai}} = \sum_a F_i$$

oznacza zerowanie się sumy sił w układzie (prawo akcji i reakcji).

Możemy teraz zdefiniować nową wielkość, którą nazywamy pędem odpowiadającym określonej współrzędnej układu. We współrzędnych kartezjańskich

$$p_i = \frac{\partial L}{\partial \dot{r}_i} = m \dot{r}_i$$

(Uwaga. Pęd pojedynczej cząstki nie jest całką ruchu. Całką ruchu jest pęd całkowity.)

Podobną wielkość możemy zdefiniować we współrzędnych uogólnionych:

$$p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}$$

Nazywamy ją pędem uogólnionym sprzężonym ze współrzędną uogólnioną q_α .

Dla układu cząstek we współrzędnych kartezjańskich pęd uogólniony odpowiadający położeniu cząstki pokrywa się ze znanym z fizyki ogólnej wyrażeniem

$$p_i = m \dot{r}_i.$$

W przypadku współrzędnych krzywoliniowych, pędy uogólnione nie muszą mieć nawet wymiaru pędu, jednakże zawsze są liniowymi funkcjami prędkości uogólnionych.

Podobnie możemy zdefiniować siłę uogólnioną

$$F_\alpha = \frac{\partial L}{\partial q_\alpha}$$

i napisać wynikający z równań Lagrange'a związek

$$\dot{p}_\alpha = F_\alpha.$$

Moment pędu.

Znajdźmy teraz całkę ruchu wynikającą z izotropowości przestrzeni. W tym celu zapiszmy wektory wodzące cząstek we współrzędnych sferycznych i dokonajmy obrotu o jednakowy dla wszystkich cząstek nieskończenie mały kąt $\delta\varphi$. Przy takim obrocie bezwzględna wartość zmiany wektora wodzącego dana jest wyrażeniem:

$$|\delta\vec{r}| = r \sin\vartheta \delta\varphi$$

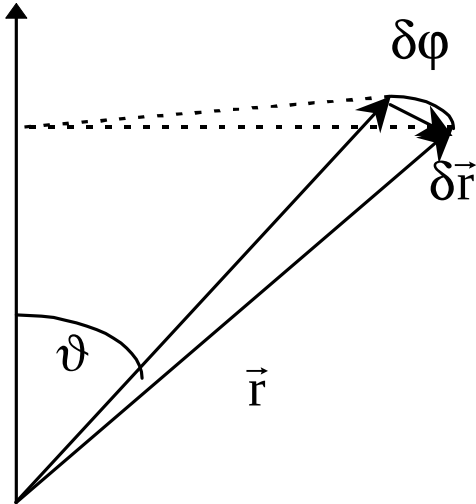
Ponieważ ponadto kierunek $\delta\vec{r}$ jest prostopadły do $\delta\vec{\varphi}$ i $\vec{r}$:

$$\delta\vec{r} = \delta\vec{\varphi} \times \vec{r}$$

a jego pochodna po czasie:

$$\delta \dot{\vec{r}} = \delta \vec{\phi} \times \dot{\vec{r}}$$

Założmy, że funkcja Lagrange'a nie zależy w sposób jawny od czasu (układ izolowany) i zażądajmy jej niezmienniczości przy obrocie $\delta \vec{\phi}$.



$$\delta L = \sum_{ai} \left(\frac{\partial L}{\partial \vec{r}_{ai}} \delta \vec{r}_{ai} + \frac{\partial L}{\partial \dot{\vec{r}}_{ai}} \delta \dot{\vec{r}}_{ai} \right) = 0$$

Korzystamy z definicji pędu i jego czasowej pochodnej:

$$\vec{p}_i = \frac{\partial L}{\partial \dot{\vec{r}}_i}, \quad \dot{\vec{p}}_i = \frac{d}{dt} \frac{\partial L}{\partial \dot{\vec{r}}_i} = \frac{\partial L}{\partial \vec{r}_i}$$

uzyskujemy:

$$\sum_{ai} (\dot{\vec{p}}_{ai} \delta \vec{r}_{ai} + \vec{p}_{ai} \delta \dot{\vec{r}}_{ai}) = 0,$$

a po wstawieniu wyrażeń uzyskanych na $\delta \vec{\phi}$ i $\delta \vec{r}$, otrzymujemy:

$$\sum_{ai} (\dot{\vec{p}}_{ai} (\delta \vec{\phi} \times \vec{r}_{ai})_i + \vec{p}_{ai} (\delta \vec{\phi} \times \dot{\vec{r}}_{ai})_i) = 0$$

lub w zapisie wektorowym:

$$\sum_a (\dot{\vec{p}}_a \cdot (\delta \vec{\phi} \times \vec{r}_a) + \vec{p}_a \cdot (\delta \vec{\phi} \times \dot{\vec{r}}_a)) = 0$$

Po zmianie kolejności czynników w iloczynie

mieszanym możemy wyciągnąć przed nawias $\delta \vec{\phi}$:

$$\delta \vec{\phi} \cdot \sum_a (\vec{r}_a \times \dot{\vec{p}}_a + \dot{\vec{r}}_a \times \vec{p}_a) = \delta \vec{\phi} \cdot \frac{d}{dt} \sum_a \vec{r}_a \times \vec{p}_a = 0$$

Korzystając z dowolności $\delta \vec{\phi}$ otrzymujemy:

$$\frac{d}{dt} \sum_a \vec{r}_a \times \vec{p}_a = 0$$

czyli uzyskaliśmy całki ruchu

$$\vec{J} = \sum_a \vec{r}_a \times \vec{p}_a = \text{const}(t)$$

Jest to wektor całkowitego momentu pędu układu. Możemy go potraktować jako sumę momentów pędu cząstek tworzących układ:

$$\vec{J} = \sum_a \vec{j}_a \quad \text{gdzie} \quad \vec{j}_a = \vec{r}_a \times \vec{p}_a$$

Całkowity moment pędu jest zachowany w przypadku układu izolowanego.

Jest zachowany również wtedy gdy w przestrzeni nie jest wyróżniony żaden kierunek, a więc na przykład dla cząstki w polu sił centralnych, tzn. gdy potencjał zależy wyłącznie od odległości od początku układu, a nie zależy od kątów.

Zmienne cykliczne a całki ruchu.

Wracając do równań Lagrange'a możemy zauważyć inny sposób poszukiwania całek ruchu.

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = 0$$

Jeżeli funkcja Lagrange'a nie zależy od pewnej współrzędnej uogólnionej np. q_α , współrzędną taką nazywamy zmienną cykliczną. Zauważmy, że wówczas:

$$\frac{\partial L}{\partial q_\alpha} = 0 = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = \frac{d}{dt} p_\alpha$$

co oznacza, że pęd uogólniony sprzężony ze zmienną cykliczną jest całką ruchu. Zilustrujmy tę możliwość poszukiwania całek ruchu prostym przykładem. Napiszmy we współrzędnych sferycznych funkcję Lagrange'a dla cząstki w polu sił centralnych:

$$L = \frac{m}{2} (\dot{r}^2 + \dot{\phi}^2 r^2 \sin^2 \vartheta + \dot{\vartheta}^2 r^2) - U(r)$$

Zauważmy, że funkcja ta nie zależy od kąta ϕ . Zatem ϕ jest zmienną cykliczną. Całką ruchu powinien być pęd sprzężony do tej zmiennej.

$$p_{\phi} = \frac{\partial L}{\partial \dot{\phi}} = m \dot{\phi} r^2 \sin^2 \vartheta$$

Łatwo sprawdzić, że pęd ten stanowi poznaną już całkę ruchu, mianowicie "zetową" składową momentu pędu.

$$\vec{J} = \vec{r} \times \vec{p}$$

Równania kanoniczne Hamiltona.

W poznanym przez nas formalizmie Lagrange'a układ mechaniczny opisywany był przy użyciu współrzędnych i prędkości uogólnionych. W pewnych zastosowaniach, wygodniej jest opisać układ przy pomocy innych wielkości fizycznych. W mechanice kwantowej nie operuje się pojęciem prędkości cząstki, gdyż ona jest naogół nieokreślona, ale pędem mającym znacznie lepszą interpretację fizyczną. Przejdźmy więc do formalizmu, w którym układ zostanie opisany przy użyciu współrzędnych i pędów uogólnionych. Nosi on nazwę formalizmu Hamiltona.

Różniczkę funkcji Lagrange'a

$$dL = \sum_{\alpha} \left(\frac{\partial L}{\partial q_{\alpha}} dq_{\alpha} + \frac{\partial L}{\partial \dot{q}_{\alpha}} d\dot{q}_{\alpha} \right)$$

korzystając z definicji pędu uogólnionego oraz z równań Lagrange'a

$$p_{\alpha} = \frac{\partial L}{\partial \dot{q}_{\alpha}}$$

$$\dot{p}_{\alpha} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_{\alpha}} = \frac{\partial L}{\partial q_{\alpha}}$$

możemy przekształcić do postaci:

$$\begin{aligned} dL &= \sum_{\alpha} (\dot{p}_{\alpha} dq_{\alpha} + p_{\alpha} d\dot{q}_{\alpha}) = \sum_{\alpha} (\dot{p}_{\alpha} dq_{\alpha} + p_{\alpha} d\dot{q}_{\alpha} + \dot{q}_{\alpha} dp_{\alpha} - \dot{q}_{\alpha} dp_{\alpha}) = \\ &= d \sum_{\alpha} p_{\alpha} \dot{q}_{\alpha} + \sum_{\alpha} (\dot{p}_{\alpha} dq_{\alpha} - \dot{q}_{\alpha} dp_{\alpha}) \end{aligned}$$

co można zapisać jako:

$$d \sum_{\alpha} (p_{\alpha} \dot{q}_{\alpha} - L) = - \sum_{\alpha} (\dot{p}_{\alpha} dq_{\alpha} - \dot{q}_{\alpha} dp_{\alpha})$$

Lewą stronę potraktujmy jako różniczkę zupełną nowej funkcji, którą nazwiemy funkcją Hamiltona:

$$H = \sum_{\alpha} (p_{\alpha} \dot{q}_{\alpha} - L) = \sum_{\alpha} \left(\frac{\partial L}{\partial \dot{q}_{\alpha}} \dot{q}_{\alpha} - L \right) = E$$

Zauważmy, że funkcja Hamiltona odpowiada zdefiniowanej poprzednio energii całkowitej układu.

Różniczka zupełna funkcji Hamiltona dana jest wyrażeniem

$$dH = \sum_{\alpha} (\dot{q}_{\alpha} dp_{\alpha} - \dot{p}_{\alpha} dq_{\alpha})$$

Możemy zauważyć bardzo istotną w tym momencie własność funkcji Hamiltona. Zgodnie z definicją różniczki zupełnej, różniczki po prawej stronie określają niezależne argumenty funkcji. Są nimi współrzędne i pędy uogólnione. Funkcja Hamiltona jest funkcją współrzędnych i pędów (pamiętamy, że funkcja Lagrange'a była funkcją współrzędnych i prędkości). Ponieważ

$$dH = \sum_{\alpha} \left(\frac{\partial H}{\partial p_{\alpha}} dp_{\alpha} + \frac{\partial H}{\partial q_{\alpha}} dq_{\alpha} \right)$$

porównując to wyrażenie z poprzednim uzyskujemy równania:

$$\frac{\partial H}{\partial p_{\alpha}} = \dot{q}_{\alpha} \quad \text{oraz} \quad \frac{\partial H}{\partial q_{\alpha}} = -\dot{p}_{\alpha}$$

Noszą one nazwę równań Hamiltona. Jest to układ 2f równań różniczkowych pierwszego rzędu. Zauważmy, że jeżeli równania Hamiltona są spełnione, wówczas:

$$\frac{dH}{dt} = \sum_{\alpha} \left(\frac{\partial H}{\partial q_{\alpha}} \dot{q}_{\alpha} + \frac{\partial H}{\partial p_{\alpha}} \dot{p}_{\alpha} \right) + \frac{\partial H}{\partial t} = \sum_{\alpha} (-\dot{p}_{\alpha} \dot{q}_{\alpha} + \dot{q}_{\alpha} \dot{p}_{\alpha}) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}$$

W przypadku braku jawnej zależności hamiltonianu od czasu otrzymujemy prawo zachowania energii.

Podsumujmy uzyskany wynik.

Funkcja Hamiltona odpowiada wyrażeniu na energię całkowitą będącą sumą energii kinetycznej i potencjalnej:

$$H = T + U,$$

w którym zależność od prędkości została zastąpiona zależnością od pędów:

$$H = H(q_{\alpha}, p_{\alpha}, t)$$

Równania Hamiltona:

$$\frac{\partial H}{\partial p_{\alpha}} = \dot{q}_{\alpha} \quad \text{oraz} \quad \frac{\partial H}{\partial q_{\alpha}} = -\dot{p}_{\alpha}$$

podobnie jak równania Lagrange'a prowadzą do równań ruchu.

Nawiasy Poissona.

W formalizmie hamiltonowskim wielkości fizyczne wyrażamy funkcjami położeń i sprzężonych z nimi pędów uogólnionych. Dotyczy to zarówno hamiltonianu określającego energię całkowitą układu jak i innych wielkości fizycznych. Rozważmy pewną wielkość fizyczną, której zależność od współrzędnych, pędów i czasu dana jest pewną funkcją $f(q, p, t)$. Zupełną pochodną czasową tej funkcji

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{\alpha} \left(\frac{\partial f}{\partial q_{\alpha}} \dot{q}_{\alpha} + \frac{\partial f}{\partial p_{\alpha}} \dot{p}_{\alpha} \right)$$

możemy przekształcić korzystając z równań Hamiltona:

$$\frac{\partial H}{\partial p_{\alpha}} = \dot{q}_{\alpha} \quad \text{oraz} \quad \frac{\partial H}{\partial q_{\alpha}} = -\dot{p}_{\alpha}$$

do postaci:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{\alpha} \left(\frac{\partial f}{\partial q_{\alpha}} \frac{\partial H}{\partial p_{\alpha}} - \frac{\partial H}{\partial q_{\alpha}} \frac{\partial f}{\partial p_{\alpha}} \right) = (f, H) + \frac{\partial f}{\partial t}$$

Występuje tutaj wyrażenie zbudowane z pochodnych cząstkowych po pędzie i położeniu wyjściowej funkcji f , oraz hamiltonianu. Wyrażenie to nosi nazwę nawiasu Poissona funkcji f i funkcji H . Ponieważ nawias Poissona wielkości fizycznej f i hamiltonianu określa jej zupełną pochodną czasową może być pomocny w poszukiwaniu całek ruchu. Jeżeli funkcja f

nie zależy jawnie od czasu, co oznacza zerowanie się jej pochodnej cząstkowej, to warunkiem by była całką ruchu jest znikanie jej nawiasu Poissona z hamiltonianem:

$$(H, f) = 0.$$

Nawiasy Poissona można zdefiniować dla dowolnej pary wielkości fizycznych zapisanych w formalizmie Hamiltona (zatem będącym funkcjami położeń i pędów). Dowolnej parze $A(q, p)$ i $B(q, p)$ przyporządkujemy funkcję noszącą nazwę nawiasów Poissona, zdefiniowaną jako:

$$(A, B) = \sum_{\alpha=1}^n \left(\frac{\partial A}{\partial q_{\alpha}} \frac{\partial B}{\partial p_{\alpha}} - \frac{\partial B}{\partial q_{\alpha}} \frac{\partial A}{\partial p_{\alpha}} \right)$$

W dalszym ciągu wykładu nie będziemy wykorzystywać nawiasów Poissona, przejdziemy bowiem do innych zagadnień, jednakże musimy im poświęcić nieco uwagi ze względu na ich rolę w przejściu od mechaniki klasycznej do kwantowej. Na czym ona polega dowiemy się w przyszłym semestrze, dzisiaj ograniczymy się do poznania kilku ich własności, do których odwołam się w przyszłym semestrze podczas wykładu mechaniki kwantowej.

1. $(A, B) = -(B, A)$,
2. $(A, A) = 0$
3. $(A + B, C) = (A, C) + (B, C)$,
4. $(AB, C) = A(B, C) + (A, C)B$,
5. $((A, B), C) + ((C, A), B) + ((B, C), A) = 0$,
6. $(A, W) = 0$, $(WA, B) = W(A, B)$ dla stałej (niezależnej od położeń i pędów) W :
7. $(q_i, q_j) = 0$, $(p_i, p_j) = 0$, $(q_i, p_j) = \delta_{ij}$
8. $(A, q_i) = -\frac{\partial A}{\partial p_i}$
9. $(A, p_i) = \frac{\partial A}{\partial q_i}$
10. $(j_x, j_y) = -j_z$

Pierwsza własność (antysymetria) jest widoczna na pierwszy rzut oka bez wykonywania żadnych rachunków:

$$(A, B) = -(B, A)$$

Z niej z kolei wynika bezpośrednio:

$$(A, A) = 0.$$

Ze względu na liniowość nawiasów Poissona względem obu wielkości fizycznych również bez trudu uzyskujemy ich rozdzielność względem sumowania:

$$(A + B, C) = (A, C) + (B, C).$$

Własność rozdzielności dla iloczynu wielkości fizycznych musimy policzyć:

$$\begin{aligned} (AB, C) &= \sum_{\alpha=1}^n \left(\frac{\partial(AB)}{\partial q_{\alpha}} \frac{\partial C}{\partial p_{\alpha}} - \frac{\partial C}{\partial q_{\alpha}} \frac{\partial(AB)}{\partial p_{\alpha}} \right) = \\ &= \sum_{\alpha=1}^n \left(B \frac{\partial A}{\partial q_{\alpha}} \frac{\partial C}{\partial p_{\alpha}} + A \frac{\partial B}{\partial q_{\alpha}} \frac{\partial C}{\partial p_{\alpha}} - B \frac{\partial C}{\partial q_{\alpha}} \frac{\partial A}{\partial p_{\alpha}} - A \frac{\partial C}{\partial q_{\alpha}} \frac{\partial B}{\partial p_{\alpha}} \right) = \\ &= \sum_{\alpha=1}^n \left(B \frac{\partial A}{\partial q_{\alpha}} \frac{\partial C}{\partial p_{\alpha}} - B \frac{\partial C}{\partial q_{\alpha}} \frac{\partial A}{\partial p_{\alpha}} \right) + \sum_{\alpha=1}^n \left(A \frac{\partial B}{\partial q_{\alpha}} \frac{\partial C}{\partial p_{\alpha}} - A \frac{\partial C}{\partial q_{\alpha}} \frac{\partial B}{\partial p_{\alpha}} \right) = \\ &= A(B, C) + B(A, C) \end{aligned}$$

Podobnie liczymy własność 5. Własność 6 wynika bezpośrednio z definicji nawiasów Poissona:

$$(A, W) = 0, (WA, B) = W(A, B),$$

a własności 7-9 również sprawdzamy w pamięci.

Podsumowanie IV.

- Wprowadziliśmy pojęcie całki ruchu jako wielkości fizycznej niezmienniej w czasie. Z ogólnych rozważań wynika, że układ o f stopniach swobody może mieć $2f-1$ niezależnych całek ruchu.
- Kilka najważniejszych całek ruchu wynika z podstawowych własności przestrzeni i czasu. Dzięki jednorodności czasu całkowitą energią jest energia, dzięki jednorodności przestrzeni pęd całkowity a całkowitą energią wynikającą z izotropowości przestrzeni jest całkowity moment pędu.
- Następnie wprowadziliśmy pojęcie zmiennej cyklicznej i pokazaliśmy, że pęd uogólniony sprzężony do zmiennej cyklicznej jest również całkowitą energią.
- Korzystając z wprowadzonej uprzednio definicji pędu uogólnionego przeszliśmy do nowego formalizmu, w którym jako zmienne niezależne traktowane są nie współrzędne i prędkości ale współrzędne i pędy. Nosi on nazwę formalizmu Hamiltona.
- W formalizmie tym rolę funkcji Lagrange'a przejmują funkcje Hamiltona, mająca sens energii całkowitej układu wyrażonej poprzez położenia i pędy. Równania ruchu generowane są przez tzw. równania kanoniczne Hamiltona. Dla f stopni swobody stanowią one układ $2f$ równań różniczkowych pierwszego rzędu względem czasu (równania Lagrange'a stanowiłyby układ f równań drugiego rzędu).
- W formalizmie Hamiltona definiujemy tzw. Nawiasy Poissona, które w mechanice klasycznej są pomocne przy poszukiwaniu całek ruchu a najważniejszym ich zastosowaniem jest przejście od mechaniki klasycznej do kwantowej.

Podsumowanie mechaniki klasycznej.

1. Mechanikę teoretyczną zbudowaliśmy postulując na wstępie zasadę najmniejszego działania:

Dla każdego układu mechanicznego można znaleźć funkcję współrzędnych i prędkości uogólnionych oraz czasu $L(q_\alpha, \dot{q}_\alpha, t)$ taką, że funkcjonał:

$$S[\tilde{q}_\alpha, \dot{\tilde{q}}_\alpha] = \int_{t_1}^{t_2} dt L(\tilde{q}_\alpha, \dot{\tilde{q}}_\alpha, t)$$

przyjmuje najmniejszą wartość dla ruchu rzeczywistego, tj. $\tilde{q}_\alpha(t) = q_\alpha(t)$. Funkcja L nosi nazwę funkcji Lagrange'a, a funkcjonał S nazywamy działaniem.

2. W ogólności funkcja Lagrange'a jest różnicą energii kinetycznej i potencjalnej:

$$L(\vec{r}, \dot{\vec{r}}, t) = T(\dot{\vec{r}}) - U(\vec{r})$$

3. Znając funkcję Lagrange'a możemy uzyskać równania ruchu wykorzystując równania Lagrange'a II rodzaju:

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = 0$$

4. Wprowadziliśmy pojęcie całki ruchu. Możemy znaleźć $2f-1$ niezależnych całek ruchu.

Niektóre z nich związane są z podstawowymi własnościami przestrzeni i czasu: $E, \vec{P}, \vec{J}$.

Ponadto całkowitą energią są pędy kanonicznie sprzężone do zmiennych cyklicznych oraz wielkości fizyczne, których nawias Poissona z hamiltonianem znika.

5. Przy pomocy funkcji Lagrange'a możemy zdefiniować pędy uogólnione:

$$p_{\alpha} = \frac{\partial L}{\partial \dot{q}_{\alpha}}$$

6. Traktując jako zmienne niezależne współrzędne uogólnione i pędy uogólnione możemy przejść do alternatywnego formalizmu Hamiltona, w którym równania ruchu uzyskujemy z równań Hamiltona

$$\frac{\partial H}{\partial p_{\alpha}} = \dot{q}_{\alpha} \quad \text{oraz} \quad \frac{\partial H}{\partial q_{\alpha}} = -\dot{p}_{\alpha}$$

gdzie funkcja Hamiltona stanowi wyrażenie na energię całkowitą układu wyrażoną jako funkcję położeń i pędów.

$$H(q_{\alpha}, p_{\alpha}, t) = T + U$$

7. W formalizmie Hamiltona każdej parze wielkości fizycznych można przypisać nawias Poissona:

$$(A, B) = \sum_{\alpha=1}^n \left(\frac{\partial A}{\partial q_{\alpha}} \frac{\partial B}{\partial p_{\alpha}} - \frac{\partial B}{\partial q_{\alpha}} \frac{\partial A}{\partial p_{\alpha}} \right)$$

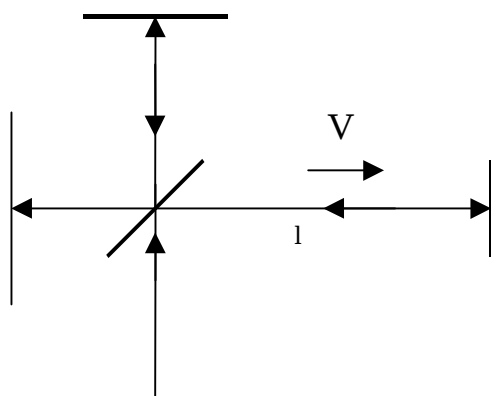
posiadający cały szereg ciekawych własności. Tych ostatnich przypominać nie będę, ponieważ omówiliśmy je przed chwilą.

(2) Mechanika relatywistyczna

V.

Zasada względności Einsteina

Mechanika klasyczna opisuje poprawnie szereg zjawisk fizycznych i pod koniec XIX w wydawało się, że przy jej pomocy można zinterpretować wszystkie zjawiska występujące w przyrodzie. Jednakże w miarę udoskonalania metod pomiarowych i rozszerzania kierunków badań eksperymentalnych zaobserwowano zjawiska, których mechanika klasyczna nie była w stanie poprawnie opisać. Pierwszym faktem doświadczalnym wykraczającym poza obszar stosowalności mechaniki klasycznej a przez to zaprzeczający jej poprawności był eksperyment Michelsona i Morleya, którzy (w 1881r) pokazali, że prędkość światła nie zależy od prędkości układu odniesienia w którym dokonujemy pomiaru.



Tymczasem zgodnie z mechaniką klasyczną, jeżeli układ porusza się z prędkością V wzdłuż długości l , to czas potrzebny na jej pokonanie przez światło byłby:

$$t_1 = \frac{l}{c-v} + \frac{l}{c+v} = \frac{l}{c^2 - v^2} (c+v + c-v) = \frac{2cl}{c^2 - v^2}.$$

Natomiast jeżeli układ obrócimy, tak by układ poruszał się prostopadle do l , potrzebny jest czas:

$$t_2 = \frac{2l}{c}. \text{ Zauważmy, że } t_1 > t_2 \text{ co nie jest zgodne z wynikiem eksperymentu.}$$

Stałość prędkości światła w układach poruszających się względem siebie jest sprzeczna z wynikiem transformacji Galileusza, której prawdziwość stanowi jedno z podstawowych założeń mechaniki klasycznej (pamiętamy, że korzystaliśmy z niej konstruując pierwszą funkcję Lagrange'a). Błędna jest zatem cała teoria klasyczna.

Problem polega na założeniu nieskończonej szybkości rozchodzenia się oddziaływań.

Założenie to jest błędne. Okazuje się mianowicie, że istnieje maksymalna prędkość przesyłania informacji, z tą maksymalną prędkością rozchodzi się światło i dlatego nazwano ją prędkością światła. Naturalnie problemy pojawiają się dopiero przy bardzo dużych prędkościach cząstek tworzących układ, dla których prędkość rozchodzenia się oddziaływań jest nie do zaniedbania.

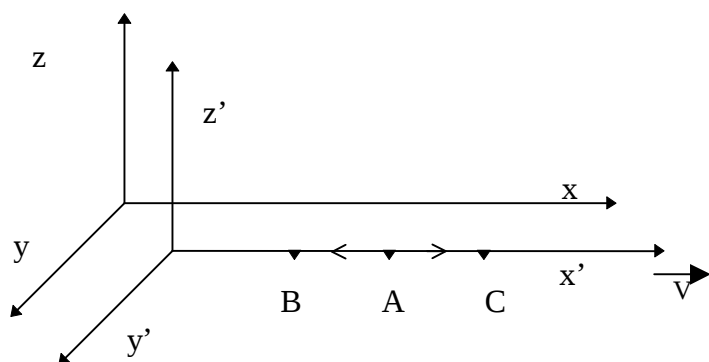
Teoria uwzględniająca nowy fakt eksperymentalny została sformułowana w 1905r. przez Einsteina. Bazuje ona na dwóch podstawowych założeniach, które noszą nazwę zasad względności Einsteina.

- Wszystkie zjawiska fizyczne przebiegają jednakowo we wszystkich inercjalnych układach odniesienia.
- Maksymalna prędkość rozchodzenia się oddziaływania (informacji) jest w każdym inercjalnym układzie odniesienia jednakowa i wynosi $c = 2.998 \cdot 10^8$ m/s.

Teoria bazująca na tej zasadzie nosi nazwę mechaniki relatywistycznej. W granicznym przypadku małych prędkości przechodzi ona w mechanikę klasyczną.

O ile pierwsza zasada nie odbiega treścią od postulowanej w mechanice klasycznej, o tyle zasada druga powoduje określone konsekwencje. Przede wszystkim zmusza do rezygnacji z bezwzględności czasu. Wyobraźmy sobie dwa układy odniesienia poruszające się względem siebie z pewną prędkością. Niech układ primowany porusza się względem nieprimowanego z prędkością V w kierunku osi x .

Z punktu A związanego z układem “primowanym” wysłane zostają sygnały w kierunku równoodległych punktów B i C związanym z tym samym układem. Sygnały obserwowane z tegoż układu osiągną oba punkty jednocześnie. Jednakże, jeżeli sygnały będą obserwowane z układu “nie primowanego”, przy założeniu że również w tym układzie poruszają się w obu



kierunkach z jednakowymi prędkościami, wówczas punkt B zostanie osiągnięty wcześniej niż punkt C , ponieważ w czasie zużyty przez sygnał punkt B przemieszcza się w kierunku przeciwnym do kierunku biegu sygnału, a punkt C odwrotnie.

Zjawisko (wysłanie sygnału świetlnego z A i jego dotarcie do punktów B i C) zachodzi w realnej przestrzeni. Układy odniesienia służą jedynie do opisu zjawiska. Tak więc pojęcie równoczesności a zatem również kolejności zdarzeń straciło bezwzględny sens i stało się zależne od układu odniesienia. Oznacza to również, że czas jako parametr numerujący kolejność zdarzeń musi zależeć od układu odniesienia.

Niezależność prędkości światła od prędkości układu odniesienia burzy pojęcie równoczesności.

Nowa teoria bazująca na zasadzie względności Einsteina jest mniej intuicyjna od mechaniki klasycznej, gdyż bezpośrednio na co dzień obserwowane zjawiska, do których jesteśmy przyzwyczajeni nie są relatywistyczne i przyzwyczailiśmy się do bezwzględności czasu. Wyniki rozważań relatywistycznych są często tak bardzo “nieoczywiste”, że należy teorię bardzo ściśle sformułować i starannie prowadzić rachunki, żeby uniknąć błędów. Wynikom rachunków, nawet tym zaskakującym można zaufać, ale pod warunkiem, że mamy pewność iż zostały poprawnie przeprowadzone.

Konstruowanie nowej teorii zaczniemy od dookreślenia pewnych nowych pojęć, którymi będziemy się posługiwać.

Zdarzenie określać będziemy przez podanie miejsca w trójwymiarowej przestrzeni i czasu. Zbiór wszystkich możliwych zdarzeń wyznacza czasoprzestrzeń.

Przedział czasoprzestrzenny.

Rozpatrzmy dwa inercjalne układy odniesienia U i U' , zakładając geometrię taką samą jak na poprzednim rysunku. Rozważmy dwa zdarzenia. Pierwsze polega na wysłaniu sygnału z punktu (x_1, y_1, z_1) w chwili t_1 , drugie na odbiorze sygnału w punkcie (x_2, y_2, z_2) w chwili t_2 .

$$(x_1, y_1, z_1, t_1) \quad \star \text{~~~~~} \star \quad (x_2, y_2, z_2, t_2)$$

Sygnał rozchodzi się z prędkością c , zatem

$$c(t_2 - t_1) = [(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]^{1/2}$$

lub

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1)^2 = 0$$

Ponieważ prędkość światła c jest jednakowa w dowolnym inercjalnym układzie odniesienia, te same dwa zdarzenia obserwowane w układzie U' dają związek:

$$(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 - c^2(t'_2 - t'_1)^2 = 0$$

Zdefiniowaną powyżej wielkość, która się zeruje dla tych dwóch zdarzeń niezależnie od układu odniesienia nazywamy przedziałem czasoprzestrzennym. Podobnie interwał (przedział) czasoprzestrzenny możemy zdefiniować dla dwóch zdarzeń dowolnej natury:

$$s_{12}^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2$$

Jeżeli oba zdarzenia są nieskończenie bliskie w przestrzeni i czasie interwał czasoprzestrzenny zapisujemy w postaci związku pomiędzy różniczkami:

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

W wyniku przyjętej definicji wiemy, że jeżeli w jednym inercjalnym układzie odniesienia przedział czasoprzestrzenny ds się zeruje, to ds' w każdym innym inercjalnym układzie odniesienia również się zeruje.

Bazując na jednorodności i izotropowości przestrzeni oraz jednorodności czasu można założyć, że wielkość ifinitezimalnego interwału czasoprzestrzennego jest jednakowa we wszystkich inercjalnych układach odniesienia:

$$ds = ds'$$

a stąd wynika równość przedziałów skończonych

$$s = s'$$

Własność ta nie wynika z definicji przedziału czasoprzestrzennego i stanowi założenie teorii prowadzącej do wybranego sposobu opisu zjawisk.

Założenie to będzie niesprzeczne z zasadami względności Einsteina, a jednocześnie pozwoli na wyprowadzenie związków pomiędzy współrzędnymi i czasem w różnych inercjalnych układach odniesienia.

Transformacja Lorentza.

Bazując na poczynionych założeniach spróbujemy przyjrzeć się zależności parametrów opisujących zdarzenie (położenie i czas) w różnych inercjalnych układach odniesienia. Rozważmy ponownie dwa układy inercjalne U i U' . Ustawmy je tak by ich odpowiednie osie były równoległe do siebie, a oś x i x' równoległe do kierunku ich wzajemnej prędkości. Niech początek układu U' porusza się w układzie U z prędkością V . Chwile początkowe dla czasów w obu układach wybierzmy tak, by odpowiadały zdarzeniu w którym położenia początków obu układów się pokrywają (wtedy zdarzenie określające początki układów w chwili początkowej mają te same parametry $(0,0,0,0)$ w obu układach). Jeżeli znajdziemy reguły przejścia pomiędzy tymi układami, będą one wystarczające do opisu przejścia pomiędzy dowolnymi układami, ponieważ wybór kierunków osi i chwil początkowych jest zawsze dowolny.

W mechanice klasycznej przejście pomiędzy różnymi inercjalnymi układami odniesienia opisywała transformacja Galileusza. Jeżeli w układzie U pewnemu zdarzeniu odpowiadają współrzędne x, y, z, t to w układzie U'

$$t' = t$$

$$x = x' + Vt$$

$$y' = y$$

$$z'=z$$

W mechanice relatywistycznej transformacja nieco się komplikuje ze względu na brak bezwzględności czasu (t' nie jest równe t).

Okazuje się na szczęście, że do wyprowadzenia związków pomiędzy współrzędnymi zdarzenia w różnych inercjalnych układach odniesienia całkowicie wystarcza wykorzystanie własności niezmienniczości przedziału czasoprzestrzennego (dlatego wprowadziliśmy to pojęcie i sprawdziliśmy związek pomiędzy jego niezmienniczością i stałością prędkości światła w różnych inercjalnych układach odniesienia).

$$s^2 = c^2 t^2 - x^2 - y^2 - z^2$$

określa przedział czasoprzestrzenny pomiędzy zdarzeniem (t, x, y, z) , a zdarzeniem określającym początek układu $(0, 0, 0, 0)$. Ponieważ układy U i U' wybraliśmy tak, by ich początki się pokrywały, odpowiedni przedział w układzie U' dany jest wyrażeniem

$$s'^2 = c^2 t'^2 - x'^2 - y'^2 - z'^2$$

Żądanie równości s i s' powinno dać nam reguły transformacyjne. Pozornie nie jest to proste, ale ułatwimy sobie sprawę, jeżeli zamiast czasu wprowadzimy nową zmienną:

$$\tau = ict$$

wówczas w układzie U

$$-s^2 = \tau^2 + x^2 + y^2 + z^2$$

i ono ma być równe

$$-s'^2 = \tau'^2 + x'^2 + y'^2 + z'^2$$

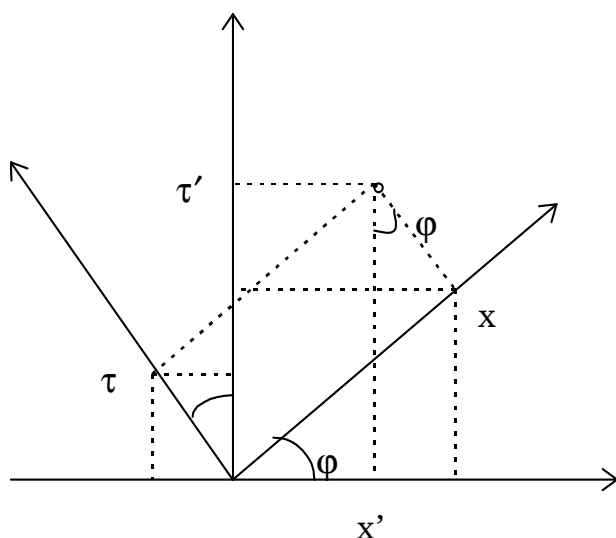
Zauważmy, że $-s^2$ oznacza kwadrat długości wektora wodzącego w przestrzeni czterowymiarowej (τ, x, y, z) . Wiemy, że długości wektorów wodzących punktów nie zmieniają się przy obrotach układu odniesienia, zatem przejścia pomiędzy układami inercjalnymi są obrotami w czterowymiarowej przestrzeni (τ, x, y, z) .

Chcemy opisać transformację pomiędzy wybranymi przez nas układami U i U' . Ponieważ założyliśmy, że oba układy poruszają się względem siebie w kierunku x (x'). Zatem współrzędne y i z nie ulegają zmianie:

$$y' = y$$

$$z' = z$$

Oznacza to obrót w płaszczyźnie (τ, x) , co znacznie ułatwia dalsze rozważania.



- Jeżeli dokonujemy obrotu o kąt ϕ , możemy wynik transformacji napisać w postaci:

$$\tau' = x \sin \phi + \tau \cos \phi$$

$$x' = x \cos \phi - \tau \sin \phi$$

Łatwo sprawdzić, że $\tau^2 + x^2 = \tau'^2 + x'^2$. Naturalnie kąt obrotu zależy może wyłącznie od prędkości względnej układu V . Żeby go wyznaczyć znajdziemy położenie początku układu U w układzie U' .

Położenie początku układu U ma w nim współrzędną

$$x=0$$

Związki transformacyjne dają jego położenie w układzie U' :

$$\tau' = \tau \cos \varphi$$

$$x' = -\tau \sin \varphi$$

τ na razie jest dowolne (warunki powinny być spełnione dla dowolnej chwili), ale możemy je wyeliminować dzieląc równania stronami

$$\frac{x'}{\tau'} = -\frac{\tau \sin \varphi}{\tau \cos \varphi} = -\operatorname{tg} \varphi$$

Ponieważ początek układu U porusza się w układzie U' z prędkością -V czyli możemy napisać związek

$$x' = -Vt'$$

lub

$$\frac{x'}{t'} = -V$$

lub

$$\frac{x'}{\tau'} = \frac{x'}{ict'} = i \frac{V}{c}$$

Porównując oba związki otrzymujemy

$$\operatorname{tg} \varphi = -i \frac{V}{c}$$

Możemy teraz znaleźć sinus i cosinus kąta φ

$$\sin \varphi = \frac{\operatorname{tg} \varphi}{\sqrt{1 + \operatorname{tg}^2 \varphi}} = \frac{-i \frac{V}{c}}{\sqrt{1 - \frac{V^2}{c^2}}}$$

$$\cos \varphi = \frac{1}{\sqrt{1 + \operatorname{tg}^2 \varphi}} = \frac{1}{\sqrt{(1 - \frac{V^2}{c^2})}}$$

Po wstawieniu uzyskujemy następujące związki:

$$\tau' = x \frac{-i \frac{V}{c}}{\sqrt{1 - \frac{V^2}{c^2}}} + \tau \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}} = \frac{\tau - i \frac{V}{c} x}{\sqrt{1 - \frac{V^2}{c^2}}}$$

$$x' = x \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}} + \tau \frac{i \frac{V}{c}}{\sqrt{1 - \frac{V^2}{c^2}}} = \frac{x + i \frac{V}{c} \tau}{\sqrt{1 - \frac{V^2}{c^2}}}$$

a po podstawieniu $\tau = ict$:

$$t' = \frac{t - \frac{V}{c^2} x}{\sqrt{1 - \frac{V^2}{c^2}}}$$

$$x' = \frac{x - Vt}{\sqrt{1 - \frac{V^2}{c^2}}}$$

$$y' = y$$

$$z' = z$$

Zauważmy, że w przypadku granicznym $V \gg c$ wzory te przechodzą we wzory transformacji Galileusza mechaniki klasycznej. Stanowi to naturalnie warunek ich poprawności.

Dla skrócenia zapisu stosuje się zwykle umowne podstawienia:

$$\beta = \frac{V}{c}$$

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

wtedy transformacja ma postać

$$t' = \gamma(t - \beta \frac{x}{c})$$

$$x' = \gamma(x - \beta ct)$$

Transformacja ta pozwala znaleźć związki pomiędzy współrzędnymi i czasem w dowolnych inercjalnych układach odniesienia. Możemy natychmiast zauważyć, że niemożliwa jest transformacja do układów poruszających się z prędkością większą od prędkości światła

$$\beta > 1 \Leftrightarrow \text{urojony}$$

gdyż wielkości fizyczne określające położenie i czas muszą być rzeczywiste.

Uwaga! W przedstawionym rachunku zastąpiliśmy t przez $\tau = ict$ w celu uzyskania pełnej analogii z obrotem w przestrzeni trójwymiarowej. Można tego nie robić, zachować zmienną czasową. Wtedy we wzorach transformacji obrotu w płaszczyźnie (x, t) należy zastąpić funkcje trygonometryczne hiperbolicznymi.

Transformacja odwrotna.

Łatwo można sprawdzić, że transformacja odwrotna odpowiada obrotowi o kąt $-\varphi$ i dana jest związkami:

$$t = \frac{t' + \frac{V}{c^2} x'}{\sqrt{1 - \frac{V^2}{c^2}}}$$

$$x = \frac{x' + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}}$$

$$y = y'$$

$$z = z'$$

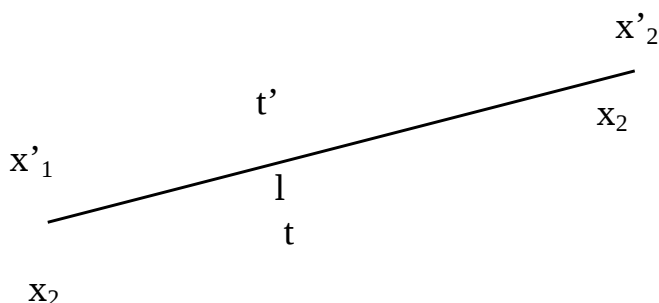
Można ją także uzyskać rozwikłując transformację prostą względem zmiennych “primowanych”.

Konsekwencją transformacji Lorentza są niespotykane w przypadku klasycznym efekty.

Przyjrzyjmy się najciekawszym.

Kontrakcja długości i dylatacja czasu.

Założmy, że w układzie U spoczywa pręt o długości l . W wybranej chwili t jednakowej (w układzie U) dla obu zdarzeń dokonujemy pomiaru położenia jego końców x_1 i x_2 . Jego długość będąca wynikiem tych pomiarów wynosi $l = x_2 - x_1$.



W układzie U' , poruszającym się względem U z prędkością V dokonujemy podobnego pomiaru wyznaczając położenia x'_1 i x'_2 w chwili t' jednakowej dla obu zdarzeń w układzie U' . Zmierzona długość jest równa $l' = x'_2 - x'_1$. Zauważmy jednak, że to co było jednocześnie w układzie U' nie było jednocześnie w układzie U . Oba zdarzenia 1 i 2 (scharakteryzowane przez położenie i czas) nie mogły być tożsame w obu układach odniesienia.

Skoncentrujmy się na pomiarze długości pręta w układzie U' . Zdarzeniom 1 i 2 odpowiada ten sam czas $t'_1 = t'_2 = t'$. Oczywiście jeżeli $t'_1 = t'_2$ to $t_1 \neq t_2$, bo

$$t_1 = \gamma \left(t'_1 + \beta \frac{x'_1}{c} \right)$$

$$t_2 = \gamma \left(t'_2 + \beta \frac{x'_2}{c} \right)$$

a x'_1 nie jest równe x'_2 (zakładamy niezerową długość pręta). Zbadajmy skutki tego faktu.

Położenia końców pręta w układzie U' wyrażamy przez współrzędne zdarzenia w układzie U :

$$x'_2 = \gamma(x_2 - \beta ct_2)$$

$$x'_1 = \gamma(x_1 - \beta ct_1)$$

Długość pręta w układzie U'

$$l' = x'_2 - x'_1 = \gamma(x_2 - x_1 - \beta c(t_2 - t_1))$$

Ponieważ chwile t_2 i t_1 nie są w układzie U tożsame, wyrażamy je ponownie przez parametry zdarzeń w układzie U' :

$$l' = \gamma \left(x_2 - x_1 - \beta c \gamma \left(t'_2 + \beta \frac{x'_2}{c} - t'_1 - \beta \frac{x'_1}{c} \right) \right)$$

Teraz korzystamy z jednoczesności pomiarów w układzie U' i uzyskujemy :

$$l' = \gamma l - \beta^2 \gamma^2 l'$$

Przekształcamy to równanie wydobywając z niego l' :

$$(1 + \beta^2 \gamma^2) l' = \gamma l$$

$$\left(1 + \frac{\beta^2}{1 - \beta^2} \right) l' = \gamma l$$

$$\frac{1}{1 - \beta^2} l' = \gamma l$$

$$\gamma^2 l' = \gamma l$$

$$l' = \frac{1}{\gamma} l = \sqrt{1 - \frac{V^2}{c^2}} l < l$$

długość pręta jest największa w układzie, w którym spoczywa. Nazywamy go układem własnym pręta (ciała). Długość ciała zmierzona w każdym innym układzie odniesienia jest mniejsza.

Jeżeli zmianie ulega długość odcinka w przestrzeni, zastanówmy się jak zachowuje się długość odcinka czasu. Niech teraz w początku układu U spoczywa zegar. Porównujemy jego wskazania ze wskazaniem zegarów poruszającego się względem niego układu U'. Ponieważ jego współrzędna przestrzenna w układzie U jest równa zero ($x=0$):

$$t' = \gamma t = \frac{t}{\sqrt{1 - \frac{V^2}{c^2}}}$$

odcinek czasu mierzony w poruszającym się układzie jest dłuższy. Najkrótszy odcinek czasu pokazują zegary w układzie w którym spoczywają. Ten czas nazywamy czasem własnym układu.

Podsumowanie V.

- Eksperymentalny fakt stałej prędkości światła w inercjalnych układach odniesienia zmusił nas do rewizji całej mechaniki klasycznej i podjęcia dyskusji całkiem od nowa. Straciliśmy uniwersalne pojęcie jednoczesności. Czas uzależniliśmy od układu odniesienia.
- Podstawą nowej teorii jest zasada względności Einsteina, w której $c = \text{const}$ jest dodatkowym założeniem.
- Wprowadziliśmy pojęcie przedziału czasoprzestrzennego $s^2 = c^2 t^2 - \vec{x}^2$.
- Zerowanie się s w jednym układzie inercjalnym implikuje zerowanie się w drugim $s = 0 \Rightarrow s' = 0$.
- Założyliśmy niezmienniczość infinitezymalnych przedziałów czasoprzestrzennych $ds = ds'$.
- Uogólniliśmy ją na przedziały skończone $s = s'$.
- Z niezmienniczości s uzyskaliśmy transformację Lorentza (szczególną).

$$t' = \gamma \left(t - \beta \frac{x}{c} \right)$$

$$x' = \gamma (x - \beta c t)$$

$$y' = y$$

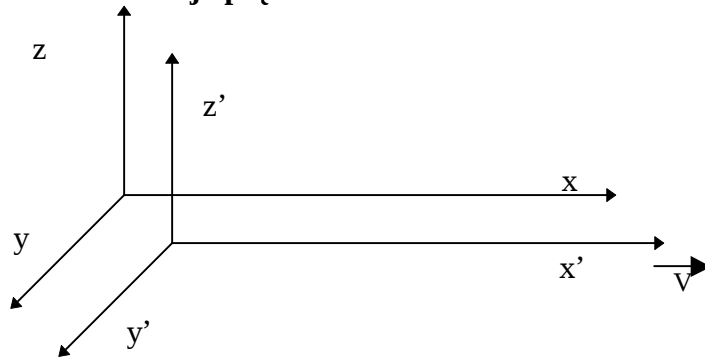
$$z' = z$$

gdzie:

$$\gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}, \beta = \frac{V}{c}$$

VI.

Transformacja prędkości.



Założmy, że w układzie U' (poruszającym się względem U z prędkością $(V, 0, 0)$) ciało posiada prędkość $\vec{v}' = (\dot{x}', \dot{y}', \dot{z}') = (\frac{dx'}{dt'}, \frac{dy'}{dt'}, \frac{dz'}{dt'})$. Znajdźmy jego prędkość względem układu U .

Klasycznie otrzymalibyśmy naturalnie $\vec{v} = \vec{v}' + \vec{V}$. W przypadku relatywistycznym musimy wykonać rachunek. Dla kierunku x :

$$\dot{x} = \frac{dx}{dt}$$

$$dx = \gamma(dx' + Vdt')$$

$$dt = \gamma(dt' + \frac{V}{c^2}dx')$$

dzielimy stronami i przekształcamy

$$\dot{x} = \frac{dx}{dt} = \frac{\gamma(dx' + Vdt')}{\gamma(dt' + \frac{V}{c^2}dx')} = \frac{\frac{dx'}{dt'} + V}{1 + \frac{V}{c^2} \frac{dx'}{dt'}} = \frac{\dot{x}' + V}{1 + \frac{V\dot{x}'}{c^2}}$$

Dla kierunków prostopadłych do $\vec{V}$:

$$\dot{y} = \frac{dy}{dt} = \frac{dy'}{\gamma(dt' + \frac{V}{c^2}dx')} = \frac{1}{\gamma} \frac{\frac{dy'}{dt'}}{1 + \frac{V}{c^2} \frac{dx'}{dt'}} = \frac{\dot{y}'}{1 + \frac{V\dot{x}'}{c^2}} \sqrt{1 - \frac{V^2}{c^2}}$$

oraz

$$\dot{z} = \frac{\dot{z}'}{1 + \frac{V\dot{x}'}{c^2}} \sqrt{1 - \frac{V^2}{c^2}}$$

Wyniki są istotnie różne niż w przypadku klasycznym. Jednakże w granicy $V \ll c$, przechodzą w klasyczny odpowiednik.

Przeanalizujmy dokładniej złożenie prędkości w kierunku x .

$$v = \frac{v' + V}{1 + \frac{Vv'}{c^2}}$$

Pokażmy, że jeżeli składane prędkości V i $v' < c$, wówczas również $v < c$.

W tym celu zróżniczkujemy v względem v'

$$\frac{dv}{dv'} = \frac{1}{1 + \frac{Vv'}{c^2}} - \frac{(v' + V)\frac{V}{c^2}}{\left(1 + \frac{Vv'}{c^2}\right)^2} = \frac{1 + \frac{Vv'}{c^2} - \frac{Vv'}{c^2} - \frac{V^2}{c^2}}{\left(1 + \frac{Vv'}{c^2}\right)^2} = \frac{1 - \frac{V^2}{c^2}}{\left(1 + \frac{Vv'}{c^2}\right)^2}$$

Ponieważ $V^2 < c^2$ uzyskane wyrażenie musi być dodatnie, v jest rosnącą funkcją v' .

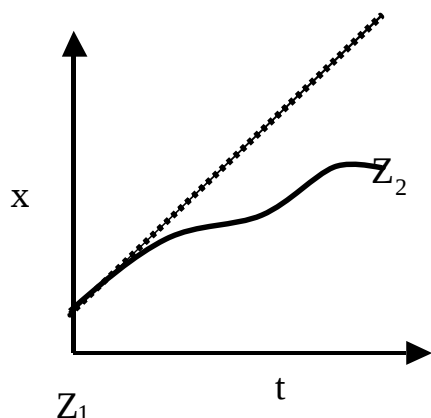
Zatem v osiąga największą wartość dla maksymalnego v' czyli dla $v' = c$ (foton w układzie U'), a wtedy

$$v = \frac{c + V}{1 + \frac{Vc}{c^2}} = c$$

Czyli maksymalna prędkość (prędkość fotonu) w układzie U jest również równa c . Tak więc transformacja Lorentza jest spójna z założoną na wstępie zasadą względności Einsteina.

Relatywistyczna całka działania.

Przejdźmy teraz do znalezienia równań ruchu w mechanice relatywistycznej. Posłużymy się zasadą najmniejszego działania i skorzystamy z formalizmu Lagrange'a podobnie jak w mechanice klasycznej. Określmy działanie dla cząstki swobodnej. Działanie powinno być niezależne od wyboru inercjalnego układu odniesienia. Poznaliśmy już pewną wielkość posiadającą tę własność. Jest nim przedział czasoprzestrzenny. Z niezmienniczości interwału czasoprzestrzennego przy przejściach pomiędzy inercjalnymi układami odniesienia korzystaliśmy przy wyprowadzaniu wzorów na transformację Lorentza. Teraz ta wielkość fizyczna posłuży nam do konstrukcji całki działania. Przyjmijmy nieskończenie mały przedział czasoprzestrzenny jako elementarne działanie. Wówczas całka interwału



czasoprzestrzennego wykonana po linii świata (czyli zbiorze kolejnych zdarzeń) od pewnego zdarzenia początkowego Z_1 do zdarzenia końcowego Z_2 stanowi działanie (z dokładnością do pewnej multiplikatywnej stałej a , którą dobierzemy później):

$$S = -a \int_{Z_1}^{Z_2} ds = -a \int_{Z_1}^{Z_2} \sqrt{c^2 dt^2 - dx^2 - dy^2 - dz^2}$$

Znak minus przed całką wybrano po to by przy dodatniej stałej a działanie posiadało minimum dla ruchu rzeczywistego. Całkę po linii świata można przekształcić w całkę po czasie

$$S = -a \int_{t_1}^{t_2} \sqrt{c^2 \frac{dt^2}{dt^2} - \frac{dx^2}{dt^2} - \frac{dy^2}{dt^2} - \frac{dz^2}{dt^2}} dt = -a \int_{t_1}^{t_2} \sqrt{c^2 - \dot{x}^2 - \dot{y}^2 - \dot{z}^2} dt =$$

$$= -ac \int_{t_1}^{t_2} \sqrt{1 - \frac{v^2}{c^2}} dt$$

i zapisać podobnie jak poprzednio jako całkę z funkcji Lagrange'a

$$S = \int_{t_1}^{t_2} L dt$$

gdzie

$$L = -ac \sqrt{1 - \frac{v^2}{c^2}}$$

Z warunku przejścia do granicy klasycznej przy $v \ll c$ wyznaczamy stałą a :
 $dla c \rightarrow \infty$

$$L = -ac \left(1 - \frac{1}{2} \frac{v^2}{c^2}\right) = -ac + \frac{av^2}{2c}$$

Odrzucając stałą $-ac$ jako zupełną pochodną czasową funkcji $-act$, dostajemy klasyczną funkcję Lagrange'a jeżeli przyjmiemy $a = m_0 c$ (m_0 oznacza klasyczną masę cząstki, wprowadzamy indeks 0 w celu odróżnienia od tzw. masy relatywistycznej, która zostanie zdefiniowana później, ale dla niej rezerwujemy literkę m bez indeksu). Teraz:

$$S = -m_0 c \int_{z_1}^{z_2} ds,$$

a funkcja Lagrange'a

$$L = -m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} = -m_0 c^2 \sqrt{1 - \frac{\dot{\vec{x}}^2}{c^2}}$$

Pęd i energia cząstki swobodnej.

Korzystając z funkcji Lagrange'a możemy znaleźć pęd cząstki:

$$p_\alpha = \frac{\partial L}{\partial q_\alpha}$$

którego składowe we współrzędnych kartezjańskich są następujące:

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = -m_0 c^2 \frac{-\frac{\dot{x}_i}{c^2}}{\sqrt{1 - \frac{\dot{\vec{x}}^2}{c^2}}} = \frac{m_0 v_i}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} = m_0 \gamma v_i$$

$$\vec{p} = \gamma m_0 \vec{v}$$

Znajdujemy również energię:

$$E = \sum_{\alpha} p_\alpha \dot{q}_\alpha - L$$

$$E = \vec{p} \cdot \vec{v} - L = \frac{m_0 v^2}{\sqrt{1 - \frac{v^2}{c^2}}} + m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} = \frac{m_0 v^2 + m_0 c^2 - m_0 v^2}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}}$$

czyli:

$$E = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma m_0 c^2$$

W granicy klasycznej $\gamma = 1$ dla pędu dostajemy wyrażenie klasyczne, ale energia przechodzi w:

$$E \approx m_0 c^2 + \frac{m_0 v^2}{2}$$

Wyrażenie to odpowiada klasycznej energii kinetycznej przesuniętej o stałą równą $m_0 c^2$. Wartość tę nazywamy energią spoczynkową cząstki.

Wyrażenie na pęd relatywistyczny zapisujemy często w postaci bardziej przypominającej klasyczną, wprowadzając zależną od prędkości cząstki masę relatywistyczną

$$m = m(v) = \gamma m_0 = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

m_0 będziemy dalej nazywać masą spoczynkową, a m masą relatywistyczną. Dla zerowej prędkości cząstki masa relatywistyczna jest naturalnie równa spoczynkowej.

Korzystając z wprowadzonej masy relatywistycznej wyrażenia na pęd i energię przepisujemy w postaci:

$$\vec{p} = m\vec{v}$$

$$E = mc^2$$

Sens wprowadzenie masy relatywistycznej jest nieco głębszy niż samo uproszczenie wyrażeń na energię i pęd. Masa występująca jako współczynnik w równaniach Newtona nosi nazwę „masy bezwładnej” ze względu na jej rolę w równaniach. Masa występuje również w wyrażeniu na siłę grawitacji. Tamta nosi nazwę „masy ciężkiej”. Okazuje się, że obie masy są ze sobą tożsame i w przypadku relatywistycznym ich rolę spełnia właśnie „masa relatywistyczna”.

Relatywistyczna zasada zachowania energii.

W związku z niezerową energią spoczynkową cząstki, prawo zachowania energii całkowitej zaczyna mieć nieco inną niż klasyczne wymowę. Jeżeli opisywany układ o energii E i masie spoczynkowej m_0 jest układem związanym, złożonym z kilku oddziaływujących pomiędzy sobą cząstek, wówczas na energię całego układu składają się energie spoczynkowe cząstek

$$\epsilon_0 = \sum_{n=1}^N m_{0n} c^2$$

ich energie kinetyczne

$$\epsilon_k = \sum_{n=1}^N E_n^k$$

oraz energie oddziaływania pomiędzy nimi

$$\epsilon_I = \sum_{i,j=1}^N U_{ij}$$

Całkowita energia spoczynkowa układu

$$E_0 = m_0 c^2 = \epsilon_0 + \epsilon_k + \epsilon_I$$

w zależności od energii kinetycznej i energii oddziaływań może być zarówno większa jak i mniejsza od sumy energii spoczynkowych składników. Jeżeli jest większa, układ nie jest stabilny i może się rozpaść wydzielając energię równą różnicy energii spoczynkowej całego układu i sumy energii spoczynkowych składników. Jeżeli masa spoczynkowa układu

związanego jest mniejsza od sumy mas spoczynkowych składników, wówczas układ związany jest stabilny a energia jest wydzielana przy jego syntezie.

Relatywistyczna funkcja Hamiltona.

Jeżeli w wyrażeniu na energię zastąpimy prędkości pędami uzyskamy relatywistyczną funkcję Hamiltona. Wyrażmy więc prędkość przez pęd.

$$\vec{p}^2 = \frac{m_0^2 \vec{v}^2}{1 - \frac{\vec{v}^2}{c^2}}$$

$$\frac{\vec{p}^2}{m_0^2 c^2} (c^2 - \vec{v}^2) = \vec{v}^2$$

$$\vec{v}^2 \left(1 + \frac{\vec{p}^2}{m_0^2 c^2}\right) = \frac{\vec{p}^2}{m_0^2}$$

$$\vec{v}^2 = \frac{\frac{\vec{p}^2}{m_0^2}}{\left(1 + \frac{\vec{p}^2}{m_0^2 c^2}\right)} = \frac{\vec{p}^2 c^2}{m_0^2 c^2 + \vec{p}^2}$$

Możemy teraz skorzystać z wyrażenia na energię zastępując prędkości pędami. Otrzymujemy

$$E = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{m_0 c^2}{\sqrt{1 - \frac{\vec{p}^2}{m_0^2 c^2 + \vec{p}^2}}} = \frac{m_0 c^2}{\sqrt{\frac{m_0^2 c^2}{m_0^2 c^2 + \vec{p}^2}}} = c \sqrt{m_0^2 c^2 + \vec{p}^2}$$

Czyli funkcja Hamiltona ma postać:

$$H = c \sqrt{m_0^2 c^2 + \vec{p}^2}$$

Transformacja energii i pędu, czterowektory.

Zastanówmy się teraz jak zmieniają się obie wprowadzone wielkości fizyczne (energia i pęd) przy zmianie układu odniesienia. Przypomnijmy sobie transformację Lorentza dla współrzędnych zdarzenia:

$$t' = \gamma \left(t - \beta \frac{x}{c}\right)$$

$$x' = \gamma (x - \beta c t)$$

gdzie

$$\gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}, \beta = \frac{V}{c}$$

Zauważmy, że przy zmianie układu odniesienia przestrzenna x i czasowa t współrzędna zdarzenia mieszają się ze sobą. Mnożąc t przez prędkość światła ujednicamy wymiary obu wielkości i możemy utworzyć obiekt, który nazywamy czterowektorem położenia i czasu: (ct, x, y, z) .

Przy transformacji Lorentza współrzędne tego wektora w układzie U' wyrażamy przez jego współrzędne w U .

Zauważmy, że również pęd i energia cząstki zależą od układu odniesienia, w którym są opisywane (zresztą w mechanice klasycznej również). Pomiędzy energią i składowymi pędu w różnych inercjalnych układach odniesienia można znaleźć związki podobne do związków pomiędzy współrzednymi i czasem. Postarajmy się znaleźć te związki.

Niech cząstka porusza się w układzie odniesienia U z prędkością v . Czas w układzie U wyrażmy przez jej czas własny, czyli czas w układzie poruszającym się względem U z prędkością v :

$$dt = \gamma d\tau$$

Zauważmy, że czas własny cząstki jest niezależny od układu w którym cząstkę opisujemy.

Prędkość w układzie U możemy wyrazić przy pomocy różniczki położenia i czasu własnego:

$$\vec{v} = \frac{d\vec{x}}{dt} = \frac{d\vec{x}}{\gamma d\tau}$$

Wstawmy to do wyrażenia na pęd cząstki w układzie U

$$\vec{p} = m_0 \gamma \vec{v} = m_0 \frac{d\vec{x}}{d\tau}$$

W wyrażeniu na energię zastępujemy γ przez $\frac{dt}{d\tau}$

$$E = m_0 \gamma c^2 = m_0 c^2 \frac{dt}{d\tau}$$

Ponieważ ani masa spoczynkowa ani czas własny nie zależą od wyboru inercjalnego układu odniesienia, natomiast od jego wyboru zależą $d\vec{x}$ i dt , możemy wykorzystać związki transformacji Lorentza pomiędzy (cdt, dx, dy, dz) dla jego odpowiednika $(E/c, p_x, p_y, p_z)$ przy przejściu pomiędzy inercjalnymi układami odniesienia. Ponieważ:

$$cdt' = \gamma(cdt - \beta dx)$$

$$dx' = \gamma(dx - \beta cdt)$$

$$dy' = dy$$

$$dz' = dz$$

Uzyskujemy związki pomiędzy odpowiednimi składowymi pędu i energią:

$$E' = m_0 c^2 \frac{dt'}{d\tau} = m_0 c^2 \gamma \left(\frac{dt}{d\tau} - \frac{\beta}{c} \frac{dx}{d\tau} \right) = m_0 c^2 \gamma \left(\frac{E}{m_0 c^2} - \frac{\beta}{c} \frac{p_x}{m_0} \right) = \gamma(E - \beta c p_x)$$

$$p'_x = m_0 \frac{dx'}{d\tau} = m_0 \gamma \left(\frac{dx}{d\tau} - \beta c \frac{dt}{d\tau} \right) = \gamma \left(p_x - \frac{\beta}{c} E \right)$$

Oba powyższe związki przepisujemy tak by upodobnić je do związków pomiędzy współrzednymi

$$\frac{E'}{c} = \gamma \left(\frac{E}{c} - \beta p_x \right)$$

$$p'_x = \gamma \left(p_x - \beta \frac{E}{c} \right)$$

$$p'_y = p_y$$

$$p'_z = p_z$$

Zauważmy, że przy zmianie inercjalnego układu odniesienia składowe pędu i energia mieszają się nawzajem. Cztery wielkości: trzy składowe pędu i energia transformują się podobnie jak trzy współrzedne przestrzenne i czas. Stanowią więc wektor w czterowymiarowej przestrzeni tzw. czterowektor. Poznaliśmy więc dwa czterowektory (cdt, dx, dy, dz) i $(E/c, p_x, p_y, p_z)$. Prędkość światła przy dt i E ujednolica wymiar współrzednych wektora. Uzyskane powyżej związki nie są jedynymi związkami pomiędzy wielkościami

fizycznymi zachodzącymi przy transformacji Lorentza. Można je zresztą znaleźć przy pomocy ogólnych geometrycznych rozważań, którym warto poświęcić odrobinę czasu.

Podsumowanie VI.

- Znaleźliśmy transformację prędkości, pokazaliśmy jej zgodność z zasadą Einsteina.
- Przeszliśmy do rekonstrukcji całej teorii (mechaniki relatywistycznej) bazując na zasadzie najmniejszego działania. Jako elementarne działanie przyjęliśmy elementarny przedział czasoprzestrzenny $dS = -m_0 c ds$.

- Uzyskaliśmy relatywistyczną funkcję Lagrange'a dla cząstki swobodnej

$$L = -m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}}.$$

- Znaleźliśmy relatywistyczne wyrażenia na najważniejsze całki ruchu: pęd i energię $\vec{p} = m\vec{v}$, $E = mc^2$, $m = \gamma m_0$.
- Określiliśmy relatywistyczną funkcję Hamiltona $H = c\sqrt{m_0^2 c^2 + \vec{p}^2}$.
- Pokazaliśmy, że przy zmianie inercjalnego układu odniesienia współrzędne pędu i energia mieszają się nawzajem i transformują się analogicznie do transformacji składowych położenia i czasu.
- Dwie czwórki wielkości (cdt, dx, dy, dz) i $(E/c, p_x, p_y, p_z)$ wykazujące analogiczne reguły transformacyjne nazwaliśmy czterowektorami.

VII.

Geometria czasoprzestrzeni, elementy rachunku tensorowego.

Wyprowadzona dotychczas transformacja Lorentza dotyczyła przejścia pomiędzy dwoma układami odniesienia których osie są odpowiednio równoległe i oś x wybrana w kierunku ich względnej prędkości.

W czasoprzestrzeni stanowi ona obrót wokół osi prostopadłej do płaszczyzny (x,t) . Nazywa się ją czasem szczególną transformacją Lorentza:

$$cdt' = \gamma(cdt - \beta dx)$$

$$dx' = \gamma(dx - \beta cdt)$$

$$dy' = dy$$

$$dz' = dz$$

Transformację pomiędzy układami dowolnie usytuowanymi względem siebie można złożyć z transformacji obrotu w przestrzeni trójwymiarowej, szczególnej transformacji Lorentza i kolejnego obrotu trójwymiarowego do poprzedniego wzajemnego usytuowania. W ogólnym przypadku związku pomiędzy współrzędnymi punktów w obu układach odniesienia są bardziej skomplikowane, jednakże zawsze nowe współrzędne są liniowymi kombinacjami starych współrzędnych:

$$x'^{\nu} = \sum_{\mu=0}^3 C^{\nu}_{\mu} x^{\mu}$$

Zauważmy, że współczynniki rozwinięcia możemy zapisać w postaci pochodnych odpowiednich współrzędnych primowanych po nieprimowanych:

$$C^{\nu}_{\mu} = \frac{\partial x'^{\nu}}{\partial x^{\mu}} x'^{\nu}$$

Transformację możemy zatem zapisać w postaci:

$$x'^{\nu} = \sum_{\mu=0}^3 \frac{\partial x'^{\nu}}{\partial x^{\mu}} x^{\mu}$$

Sumowanie przebiega po czterech składowych w czasoprzestrzeni. Do określania sumy po czterech składowych będziemy w dalszym ciągu używać wskaźników greckich przyjmujących wartości od 0 do 3, łacińskie (od 1 do 3) pozostawimy do określania składowych wektora w przestrzeni trójwymiarowej. Przyjmijmy ponadto następującą kolejność składowych $(x^{\mu}) = (ct, x, y, z)$, czas stanowi zerową składową czterowektora położenia.

W przypadku szczególnej transformacji Lorentza współczynniki transformacji tworzą następującą macierz:

$$(C^{\nu}_{\mu}) = \left(\frac{\partial x'^{\nu}}{\partial x^{\mu}} \right) = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Z transformacją współrzędnych przy zmianie układu odniesienia wiążą się ściśle własności obiektów geometrycznych, które nazywamy tensorami. Są one niezwykle użyteczne w rachunkach relatywistycznych, dlatego warto je wprowadzić i wykorzystać.

Skalarem (tensorem 0 rzędu) nazywamy obiekt geometryczny opisywany jedną liczbą niezmienniczy względem transformacji układu współrzędnych. Skalarem w czasoprzestrzeni jest więc masa spoczynkowa, czas własny, interwał czasoprzestrzenny, nie jest natomiast czas, energia, nie są składowe pędu.

Wektorem kontrawariantnym (tensorem 1 rzędu) w n wymiarowej przestrzeni nazywamy obiekt geometryczny opisywany przy pomocy n liczb, transformujący się zgodnie z transformacją współrzędnych. Tak więc wektorem w czasoprzestrzeni jest oczywiście czterowektor położenia i czasu, jest nim również czterowektor energii i pędu

$$(p^\mu) = \left(\frac{E}{c}, p_x, p_y, p_z \right)$$

(za p^0 przyjęliśmy E/c). Jak już się przekonaliśmy spełnia on regułę transformacyjną taką samą jak czterowektor położenia i czasu:

$$p'^\nu = \sum_{\mu=0}^3 C^\nu_\mu p^\mu = \sum_{\mu=0}^3 \frac{\partial x'^\nu}{\partial x^\mu} p^\mu$$

Wiele wielkości fizycznych ma charakter wektorowy w przestrzeni trójwymiarowej. Jeżeli uda się dopasować czwartą (zerową) współrzędną i skompletować czterowektor, wówczas mamy rozwiązany problem transformacji takiej wielkości fizycznej przy zmianie układu odniesienia.

Jednakże nie zawsze się to udaje. Nie wszystkie wielkości fizyczne daje się przedstawić jako skalary lub składowe czterowektora. Na przykład ładunek punktowy w układzie własnym generuje wyłącznie pole elektrostatyczne, w układzie względem którego się porusza generuje również pole magnetyczne. W rezultacie składowe pola elektrycznego i magnetycznego mieszają się nawzajem przy transformacji Lorentza (wkrótce się o tym przekonamy):

$$E'_z = \gamma(E_z - \beta c B_y)$$

Trzy składowe pola elektrycznego i trzy magnetycznego nie mogą stanowić współrzędnych czterowektora. Dlatego musimy posunąć się jeszcze o jeden krok dalej i zdefiniować tensory wyższych rzędów (będące w sensie własności transformacyjnych uogólnieniem wektorów). Tensorem kontrawariantnym 2-giego rzędu nazwiemy obiekt geometryczny opisywany przy pomocy n^2 liczb, których składowe transformują się odpowiednio zgodnie z transformacją współrzędnych:

$$T'^{\mu\nu} = \sum_{\mu_1\nu_1} \frac{\partial x'^\mu}{\partial x^{\mu_1}} \frac{\partial x'^\nu}{\partial x^{\nu_1}} T^{\mu_1\nu_1}$$

Wskaźniki μ i ν przebiegają w czasoprzestrzeni cztery wartości od 0 do 3, zatem $n=4$ a $n^2=16$, zatem tensor drugiego rzędu posiada 16 współrzędnych.

Liczbę niezależnych współrzędnych tensora można ograniczyć narzucając na jego współrzędne warunek symetrii lub antysymetrii. Własności te są zachowywane przy transformacji Lorentza. Tensor symetryczny spełnia warunek:

$$T^{\mu\nu} = T^{\nu\mu}$$

i posiada $\frac{n^2 - n}{2} + n = \frac{n^2 + n}{2} = 10$ niezależnych współrzędnych, a tensor antysymetryczny, spełniający warunek:

$$T^{\mu\nu} = -T^{\nu\mu}$$

(dzięki któremu składowe diagonalne się zerują) posiada $\frac{n^2 - n}{2} = 6$ niezależnych

składowych. Podobnie definiujemy tensory kontrawariantne wyższych rzędów.

Tensory kowariantne.

W rachunkach ważną rolę spełnia możliwość tworzenia wielkości skalarnych z wielkości wektorowych (niezmienniki względem transformacji, np. interwał czasoprzestrzenny konstruujemy z wektora położenia w czasoprzestrzeni). Dla ułatwienia zapisu wprowadzimy pojęcie drugiego rodzaju tensorów, mianowicie tensorów kowariantnych. Są to podobnie jak

poprzednio obiekty geometryczne, ale transformujące się zgodnie z transformacją odwrotną do transformacji układu współrzędnych

$$cdt = \gamma(cdt' + \beta dx')$$

$$dx = \gamma(dx' + \beta cdt')$$

$$dy = dy'$$

$$dz = dz'$$

Macierz takiej transformacji jest:

$$\left(\frac{\partial x^\mu}{\partial x'^\nu} \right) = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Dla odróżnienia od tensorów kontrawariantnych wskaźniki składowych tensora kowariantnego umieścimy u dołu. Zatem kowariantny wektor transformuje się

$$p'_\nu = \sum_{\mu=0}^3 \frac{\partial x^\mu}{\partial x'^\nu} p_\mu,$$

a kowariantny tensor rzędu drugiego:

$$T'_{\mu\nu} = \sum_{\mu_1\nu_1} \frac{\partial x^{\mu_1}}{\partial x'^\mu} \frac{\partial x^{\nu_1}}{\partial x'^\nu} T_{\mu_1\nu_1}.$$

Wielkości fizyczne mogą być zapisywane przy pomocy obu typów tensorów.

Przy pomocy wektorów kontra i kowariantnego można utworzyć skalar:

$$\sum_{\mu} q'^\mu p'_\mu = \sum_{\mu\nu\nu_1} \frac{\partial x'^\mu}{\partial x'^\nu} q^\nu \frac{\partial x^{\nu_1}}{\partial x'^\mu} p_{\nu_1} = \sum_{\nu\nu_1} \left(\sum_{\mu} \frac{\partial x'^\mu}{\partial x'^\nu} \frac{\partial x^{\nu_1}}{\partial x'^\mu} \right) q^\nu p_{\nu_1}$$

Zauważmy, że

$$\sum_{\mu} \frac{\partial x^{\nu_1}}{\partial x'^\mu} \frac{\partial x'^\mu}{\partial x'^\nu} = \frac{\partial x^{\nu_1}}{\partial x'^\nu} = \delta_{\nu}^{\nu_1} = \begin{cases} 1 & \text{gdy } \nu = \nu_1 \\ 0 & \text{gdy } \nu \neq \nu_1 \end{cases}$$

zatem

$$\sum_{\mu} q'^\mu p'_\mu = \sum_{\nu\nu_1} \delta_{\nu}^{\nu_1} q^\nu p_{\nu_1} = \sum_{\mu} q^\mu p_\mu$$

jest jednakowe w dowolnym inercjalnym układzie odniesienia. Stanowi zatem skalar.

Tensor metryczny.

Poznajmy przykład tensora drugiego rzędu. Napiszmy ogólne wyrażenie na kwadrat odległości dwóch punktów w przestrzeni odległych o wektor (dx^μ) . Wyraża się on formą kwadratową:

$$ds^2 = \sum_{\mu\nu} g_{\mu\nu} dx^\mu dx^\nu$$

Zauważmy, że przy zmianie układu odniesienia (obrotach w czterowymiarowej czasoprzestrzeni):

$$\begin{aligned} ds^2 &= \sum_{\mu\nu} g_{\mu\nu} dx^\mu dx^\nu = \sum_{\mu\nu\mu_1\nu_1} g_{\mu\nu} \frac{\partial x^\mu}{\partial x'^{\mu_1}} \frac{\partial x^\nu}{\partial x'^{\nu_1}} dx'^{\mu_1} dx'^{\nu_1} = \\ &= ds'^2 = \sum_{\mu\nu} g'_{\mu\nu} dx'^\mu dx'^\nu \end{aligned}$$

współczynniki $g_{\mu\nu}$ mają określoną regułę transformacji:

$$g'_{\mu_1 \nu_1} = \sum_{\mu \nu} g_{\mu \nu} \frac{\partial x^\mu}{\partial x'^{\mu_1}} \frac{\partial x^\nu}{\partial x'^{\nu_1}}$$

Ponieważ transformują się zgodnie z transformacją odwrotną do transformacji współrzędnych, $g_{\mu \nu}$ stanowi tensor kowariantny, nosi on nazwę kowariantnego tensora metrycznego.

W trójwymiarowej przestrzeni (euklidesowej), w której $ds^2 = dx^2 + dy^2 + dz^2$, tensor metryczny $g_{ij} = \delta_{ij}$. Natomiast w czasoprzestrzeni (pseudoeuklidesowej)

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

tensor metryczny ma postać

$$(g_{\mu \nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Korzystając z macierzy określającej współczynniki szczególnej transformacji Lorentza, można sprawdzić jego postać w innych inercjalnych układach odniesienia:

$$g'_{\mu \nu} = \sum_{\mu_1 \nu_1} \frac{\partial x^{\mu_1}}{\partial x'^{\mu}} \frac{\partial x^{\nu_1}}{\partial x'^{\nu}} g_{\mu_1 \nu_1}$$

Do wyliczenia współczynników tensora w nowych współrzędnych wykorzystamy zapis macierzowy pilnując kolejności sumowania po wierszach i kolumnach:

$$\begin{aligned} (g'_{\mu \nu}) &= \left(\frac{\partial x^\mu}{\partial x'^{\mu}} \right) (g_{\mu \nu}) \left(\frac{\partial x^\nu}{\partial x'^{\nu}} \right)^T = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ -\gamma\beta & -\gamma & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \gamma^2 - \gamma^2\beta^2 & \gamma^2\beta^2 - \gamma^2\beta^2 & 0 & 0 \\ \gamma^2\beta^2 - \gamma^2\beta^2 & \gamma^2\beta^2 - \gamma^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \end{aligned}$$

Tak więc jego składowe nie ulegają zmianie przy przejściu do innego inercjalnego układu odniesienia.

Tensor metryczny pozwala na zamianę tensora kontrawariantnego na kowariantny:

$h_\mu = \sum_\nu g_{\mu \nu} p^\nu$, który oznaczamy taką samą literką $h_\mu = p_\mu$, ponieważ odpowiada tej samej wielkości fizycznej.

Przykład: $(dx^\nu) = (ct, dx, dy, dz)$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} cdt \\ -dx \\ -dy \\ -dz \end{pmatrix}$$

zatem $(dx_v) = (ct, -dx, -dy, -dz)$

Kontrawariantna wersja czterowektora energii i pędu

$p^\mu = (\frac{E}{c}, p_x, p_y, p_z)$ ma kowariantny odpowiednik $p_\mu = (\frac{E}{c}, -p_x, -p_y, -p_z)$.

Podobnie możemy zdefiniować kontrawariantny tensor metryczny

$$ds^2 = \sum_{\mu\nu} g^{\mu\nu} dx_\mu dx_\nu = c^2 dt^2$$

$$ds^2 = \sum_{\mu\nu} g^{\mu\nu} dx_\mu dx_\nu = c^2 dt^2 - (-dx)^2 - (-dy)^2 - (-dz)^2$$

Zauważmy, że tensor metryczny kowariantny ma takie same składowe co tensor metryczny kontrawariantny

$$g^{\mu\nu} = g_{\mu\nu}.$$

Jest to ogólna własność tzw. płaskiej czasoprzestrzeni. W przypadku czasoprzestrzeni zakrzywionej, którą posługuje się ogólna teoria względności własność ta nie zachodzi.

Przy pomocy tensora metrycznego możemy zamieniać tensory kontrawariantne w kowariantne i odwrotnie. Oznacza to obniżanie lub podnoszenie wskaźników tensora. W tensorach wyższych rzędów (już dla drugiego) możemy podnieść część wskaźników.

Otrzymany obiekt nosi nazwę tensora mieszanego:

$$T^\mu{}_\nu = \sum_{\nu'} g_{\nu\nu'} T^{\mu\nu'}$$

Tensorowe własności operatorów różniczkowych.

Zauważmy, że operatory różniczkowania po zmiennych przestrzennych mają określone reguły transformacyjne przy zmianie układu odniesienia. Mają więc własności tensorowe.

Operator różniczkowania po kontrawariantnej zmiennej przestrzennej

$$\frac{\partial}{\partial x'^\mu} = \sum_\nu \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu}$$

jest wektorem (tensorem pierwszego rzędu) kowariantnym i oznaczamy go często jako

$$\frac{\partial}{\partial x^\mu} = \partial_\mu$$

Podobnie operator różniczkowania po składowych wektora kowariantnego ma własności wektora kontrawariantnego

$$\frac{\partial}{\partial x_\mu} = \partial^\mu,$$

a operator różniczkowy drugiego stopnia utworzony z operatorów różniczkowania po składowych kontra i kowariantnych

$$\sum_\mu \partial_\mu \partial^\mu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}$$

jest niezmienniczy względem transformacji Lorentza.

Podsumowanie VII.

- Wprowadziliśmy pojęcie tensora jako obiektu geometrycznego o określonej regule transformacji.
- Jeżeli wiemy jakie własności transformacyjne ma wielkość fizyczna (jest tensorem) możemy znając jej składowe tensorowe w jednym układzie przejść do innego układu inercyjnego.
- Poznaliśmy:
 - ♦ tensory kontrawariantne różnych rzędów: s , V^μ , $T^{\mu\nu}$ (skalar, wektor, tensor drugiego rzędu), transformujące się zgodnie z transformacją współrzędnych
 - ♦ tensory kowariantne : s , V_μ , $T_{\mu\nu}$ transformujące się zgodnie z transformacją odwrotną do transformacji współrzędnych.
- Potrafimy zamieniać tensory kontrawariantne na kowariantne przy pomocy tensora metrycznego:
$$V^\mu = \sum_\nu g^{\mu\nu} V_\nu, \quad V_\mu = \sum_\nu g_{\mu\nu} V^\nu$$
- Kombinując tensory możemy tworzyć tensory innych rzędów, a w szczególności skalary:
$$s = \sum_\mu a^\mu b_\mu = \sum_\mu a_\mu b^\mu = \sum_{\mu\nu} g_{\mu\nu} a^\mu b^\nu = \sum_{\mu\nu} g^{\mu\nu} a_\mu b_\nu$$

VIII.

(3) Elektrodynamika

Teoria Maxwella powstała kilkadziesiąt lat przed sformułowaniem szczególnej teorii względności (1854-1867). Jednakże związek obu teorii jest tak ścisły, że aby to pokazać skonstruujemy elektrodynamikę w odwróconej historycznie kolejności, startując z teorii relatywistycznej przejdziemy do sformułowania podstawowych praw elektrodynamiki. Posłużymy się konsekwentnie zasadą najmniejszego działania. Pracować będziemy w czasoprzestrzeni, granicę klasyczną uzyskamy przez przejście z prędkością światła do nieskończoności.

Ponieważ będziemy stosować zapis tensorowy, korzystnie jest wprowadzić konwencję sumowania Einsteina, w której powtarzający się wskaźnik tensorowy na różnych poziomach oznacza wykonanie po nim sumowania. Np.

$$a_\mu b^\mu \Leftrightarrow \sum_\mu a_\mu b^\mu = s$$

$$a_\mu T^{\mu\nu} \Leftrightarrow \sum_\mu a_\mu T^{\mu\nu} = V^\nu$$

$$\partial_\mu T^{\mu\nu} \Leftrightarrow \sum_\mu \partial_\mu T^{\mu\nu} = \sum_\mu \frac{\partial}{\partial x^\mu} T^{\mu\nu} = V^\nu$$

Cząstka w zewnętrznym polu elektrycznym i magnetycznym, czteropotencjał pola elektromagnetycznego.

Rozpatrzmy problem cząstki relatywistycznej poruszającej się w zewnętrznym polu elektromagnetycznym.

Działanie cząstki swobodnej zostało wprowadzone poprzednio, oznaczmy go jako:

$$S_m = -m_0 c \int_{z_1}^{z_2} ds$$

Do działania cząstki swobodnej musimy dodać wyraz opisujący oddziaływanie z polem elektromagnetycznym. Wprowadzamy tzw. czteropotencjał pola elektromagnetycznego, będący wektorem w czasoprzestrzeni.

$A^\mu = (\phi, c\vec{A})$ (wektor kontrawariantny)

Jego kowariantny odpowiednik ma składowe

$$A_\mu = (\phi, -c\vec{A})$$

Pierwsza składowa ϕ nosi nazwę skalarnego, druga $\vec{A}$ wektorowego potencjału pola elektromagnetycznego (występujące w nazwie pojęcia skalarny i wektorowy pochodzą z elektrodynamiki klasycznej i nie należy ich mylić z własnościami transformacyjnymi w czasoprzestrzeni). Zakładamy, że czteropotencjał zawiera informację o polach elektrycznym i magnetycznym, na razie nie precyzujemy w jaki sposób.

Działanie reprezentujące oddziaływanie cząstki z polem powinno zależeć liniowo od pola.

Najprostszym skalarom liniowym w A jest:

$$A_\mu dx^\mu \quad (\text{w wyrażeniu tym użyto konwencji sumowania}).$$

Przyjmijmy więc dodatkowe działanie w postaci

$$S_{mf} = -\frac{e}{c} \int_{z_1}^{z_2} A_\mu dx^\mu$$

Współczynnik i jego znak mogą być dowolne dopóki nie dookreślimy A .

Pełne działanie

$$S = S_m + S_{mf} = -m_0 c \int_{z_1}^{z_2} ds - \frac{e}{c} \int_{z_1}^{z_2} A_\mu dx^\mu = \int_{z_1}^{z_2} (-m_0 c ds - e\phi dt + e\vec{A} d\vec{r})$$

możemy podobnie jak poprzednio zamienić na całkę po czasie

$$S = \int_{t_1}^{t_2} \left(-m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} - e\phi + e\vec{A} \cdot \vec{v} \right) dt$$

Mamy więc funkcję Lagrange'a

$$L = -m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} - e\phi + e\vec{A} \cdot \vec{v}$$

Traktując pole ϕ i $\vec{A}$ jako zewnętrzne możemy wyliczyć pęd układu:

$$\vec{P} = \frac{\partial L}{\partial \vec{v}} = \frac{m_0 \vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} + e\vec{A} = \vec{p} + e\vec{A}.$$

Występujące w powyższym wyrażeniu różniczkowanie po wektorze będziemy rozumieć jako wektor o składowych będących pochodnymi po odpowiednich składowych.

$$P^i = \frac{\partial L}{\partial v^i}$$

Oznaczony dużą literą P pęd układu cząstka + pole, jest sumą oznaczonego małą literą p pędu cząstki (bez pola) i potencjału wektorowego pomnożonego przez ładunek cząstki.

Mając pęd możemy znaleźć energię

$$\begin{aligned} E = \vec{P} \cdot \vec{v} - L &= \vec{v} \vec{p} + e\vec{A} \cdot \vec{v} + m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} + e\phi - e\vec{A} \cdot \vec{v} = \\ &= \frac{m_0 v^2}{\sqrt{1 - \frac{v^2}{c^2}}} + m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} + e\phi = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}} + e\phi \end{aligned}$$

a po zastąpieniu prędkości pędami:

$$v^2 = \frac{p^2 c^2}{m_0^2 c^2 + p^2}$$

otrzymujemy funkcję Hamiltona dla cząstki w polu:

$$H = c \sqrt{m_0^2 c^2 + \vec{p}^2} + e\phi = c \sqrt{m_0^2 c^2 + (\vec{P} - e\vec{A})^2} + e\phi$$

Interpretacja czteropotencjału.

Korzystając z równań Lagrange'a uzyskajmy równania ruchu:

$$\frac{d}{dt} \frac{\partial L}{\partial \vec{v}} = \frac{\partial L}{\partial \vec{r}}$$

$$\dot{\vec{P}} = \frac{d}{dt} (\vec{p} + e\vec{A}) = -e\nabla\phi + e\nabla(\vec{A} \cdot \vec{v})$$

Ostatni wyraz przekształcamy korzystając z tożsamości (bez dowodu):

$$\nabla(\vec{a} \cdot \vec{b}) = (\vec{a} \cdot \nabla)\vec{b} + (\vec{b} \cdot \nabla)\vec{a} + \vec{b} \times \nabla \times \vec{a} + \vec{a} \times \nabla \times \vec{b}$$

Przyjmujemy $\vec{a} = \vec{A}$, $\vec{b} = \vec{v}$, zatem:

$$\nabla(\vec{A} \cdot \vec{v}) = (\vec{A} \cdot \nabla)\vec{v} + (\vec{v} \cdot \nabla)\vec{A} + \vec{v} \times \nabla \times \vec{A} + \vec{A} \times \nabla \times \vec{v}$$

Wyrazy pierwszy i czwarty zawierają pochodne składowych prędkości liczone po zmiennych przestrzennych. Ponieważ prędkości są wielkościami niezależnymi wyrazy te zerują się. Otrzymujemy:

$$\frac{d}{dt}(\vec{p} + e\vec{A}) = -e\nabla\phi + e(\vec{v} \cdot \nabla)\vec{A} + e\vec{v} \times \nabla \times \vec{A}$$

Ponieważ zupełna pochodna czasowa wektora $\vec{A}$ ma postać:

$$\frac{d}{dt}\vec{A} = \frac{\partial\vec{A}}{\partial t} + \sum_i \frac{\partial\vec{A}}{\partial x_i} \frac{\partial x_i}{\partial t} = \frac{\partial\vec{A}}{\partial t} + (\vec{v} \cdot \nabla)\vec{A}$$

$$\frac{d}{dt}\vec{p} + e\frac{\partial\vec{A}}{\partial t} + e(\vec{v} \cdot \nabla)\vec{A} = -e\nabla\phi + e(\vec{v} \cdot \nabla)\vec{A} + e\vec{v} \times \nabla \times \vec{A}$$

uzyskujemy:

$$\frac{d}{dt}\vec{p} = -e\nabla\phi - e\frac{\partial\vec{A}}{\partial t} + e\vec{v} \times \nabla \times \vec{A}$$

Dotychczas nie określiliśmy jeszcze znaczenia czteropotencjału, możemy to teraz zrobić, żądając by wyrażenie po prawej stronie było znaną z eksperymentu siłą Lorentza .

$$\frac{d\vec{p}}{dt} = \vec{F} = e\vec{E} + e\vec{v} \times \vec{B}$$

Porównując ze sobą oba wyrażenia uzyskujemy interpretację wprowadzonych potencjałów elektromagnetycznych:

$$\vec{E} = -\nabla\phi - \frac{\partial\vec{A}}{\partial t}$$

oraz

$$\vec{B} = \nabla \times \vec{A}$$

Związki te są znane z elektrodynamiki klasycznej, w której $\vec{A}$ definiuje się jako wektorowy, a ϕ jako skalarny potencjał pola elektromagnetycznego. Uzyskaliśmy więc interpretację składowych zapostulowanego na wstępie relatywistycznego czterowektora opisującego pole elektromagnetyczne. Interpretacja ta pozwala uzgodnić teorię relatywistyczną ze znaną wcześniej teorią nierelatywistyczną.

Niezmienniczość cechowania potencjałów elektromagnetycznych.

Czteropotencjał jest w elektrodynamice jedynie wielkością pomocniczą (nie jest bezpośrednio mierzalny), jednakże jest często wykorzystywany w rachunkach. Dlatego warto poświęcić trochę uwagi jego własnościom.

Okazuje się, że potencjały pola elektromagnetycznego nie są określone jednoznacznie przez pola elektryczne i magnetyczne. Możemy się o tym przekonać definiując nowy czteropotencjał różny od poprzedniego o czterowektor stanowiący "czterogradient" pewnej dowolnej funkcji określonej w czasoprzestrzeni.

$$\tilde{A}_\mu = A_\mu - \frac{\partial f}{\partial x^\mu} = A_\mu + \Delta A_\mu$$

W całe działania pojawia się zmiana:

$$\Delta S = -\frac{e}{c} \int_{z_1}^{z_2} \Delta A_\mu dx^\mu,$$

co po podstawieniu daje:

$$\Delta S = \frac{e}{c} \int_{z_1}^{z_2} \frac{\partial f}{\partial x^\mu} dx^\mu = \frac{e}{c} \int_{z_1}^{z_2} df.$$

Różniczka zupełna dzięki warunkom brzegowym daje po scałkowaniu zerowy wkład do wariacji działania. Równania ruchu (pola elektryczne i magnetyczne) liczone z nowym potencjałem nie powinny ulec zmianie. Można to sprawdzić odejmując składowe czterodivergencji

$$\frac{\partial f}{\partial x^\mu} = \partial_\mu f = \left(\frac{1}{c} \frac{\partial f}{\partial t}, \nabla f \right)$$

do składowych czteropotencjału:

$$\begin{aligned} \tilde{\varphi} &= \varphi - \frac{1}{c} \frac{\partial f}{\partial t} & A_\mu &= (\varphi, -\vec{A}c) \\ \tilde{\vec{A}} &= \vec{A} + \frac{1}{c} \nabla f & -c\tilde{\vec{A}} &= -c\vec{A} - \nabla f \end{aligned}$$

Wstawiając uzyskane wyrażenia do wzorów na wektory pola otrzymujemy:

$$\tilde{\vec{E}} = -\nabla \tilde{\varphi} - \frac{\partial \tilde{\vec{A}}}{\partial t} = -\nabla \varphi + \frac{1}{c} \nabla \frac{\partial f}{\partial t} - \frac{\partial \vec{A}}{\partial t} - \frac{1}{c} \frac{\partial}{\partial t} \nabla f = -\nabla \varphi - \frac{\partial \vec{A}}{\partial t} = \vec{E}$$

oraz

$$\tilde{\vec{B}} = \nabla \times \tilde{\vec{A}} = \nabla \times \vec{A} - \frac{1}{c} \nabla \times \nabla f = \nabla \times \vec{A} = \vec{B}$$

Niejednoznaczność wyboru potencjałów nosi nazwę niezmienniczości cechowania i może być wykorzystana do wyboru potencjałów ułatwiających rachunek. Przykładowo w przypadku gdy φ nie zależy od czasu (problem statyczny) można zażądać aby potencjał skalarny był równy zero:

$$\tilde{\varphi} = 0 \Rightarrow \frac{\partial f}{\partial t} = c\varphi \rightarrow f = ct\varphi$$

wtedy do potencjału wektorowego musimy dodać gradient $\frac{1}{c}f$, zatem:

$$\tilde{\vec{A}} = \vec{A} + \frac{1}{c} \nabla ct\varphi = \vec{A} + t\nabla \varphi$$

a wektory pola

$$\tilde{\vec{E}} = -\frac{\partial \tilde{\varphi}}{\partial t} = -\frac{\partial \vec{A}}{\partial t} - \frac{\partial}{\partial t} \nabla (t\varphi) = -\frac{\partial \vec{A}}{\partial t} - \nabla \varphi = \vec{E}$$

oraz

$$\tilde{\vec{B}} = \nabla \times \tilde{\vec{A}} = \nabla \times \vec{A} - \nabla \times \nabla t\varphi = \nabla \times \vec{A} = \vec{B}$$

nie ulegają zmianie.

Transformacja Lorentza dla pól elektrycznego i magnetycznego.

Korzystając ze związków pomiędzy wektorami pola a potencjałami elektromagnetycznymi można wykorzystując reguły transformacyjne czterowektora znaleźć wzory zmiany pól elektrycznego i magnetycznego przy przejściach pomiędzy różnymi układami inercjalnymi. Rachunek bezpośredni jest skomplikowany. Znakomite ułatwienie stanowi tzw. tensor pola elektromagnetycznego, którego składowymi są wektory pola.

$$f_{\mu\nu} \stackrel{\text{def}}{=} \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}$$

Tensor ten jest antysymetryczny $f_{\mu\nu} = -f_{\nu\mu}$, zerują się więc jego elementy diagonalne, a pozadiagonalne w górnej połowie różnią się jedynie znakiem od elementów w dolnej. Posiada więc 6 niezależnych elementów.

Tensor utworzony jest przy pomocy dwóch czterowektorów – czterowektora położenia i czasu:

$$(x^\mu) = (ct, x, y, z)$$

oraz zdefiniowanego wcześniej czteropotencjału (w postaci kowariantnej):

$$A_\mu = (\varphi, -c\vec{A})$$

Zgodnie z uzyskaną wcześniej interpretacją czteropotencjału, składowe pola elektrycznego spełniają związki:

$$\vec{E} = -\nabla\varphi - \frac{\partial\vec{A}}{\partial t}$$

a składowe pola magnetycznego:

$$\vec{B} = \nabla \times \vec{A} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = \vec{i} \left(\frac{\partial}{\partial y} A_z - \frac{\partial}{\partial z} A_y \right) + \vec{j} \left(\frac{\partial}{\partial z} A_x - \frac{\partial}{\partial x} A_z \right) + \vec{k} \left(\frac{\partial}{\partial x} A_y - \frac{\partial}{\partial y} A_x \right)$$

Korzystając z tych własności znajdujemy elementy tensora f.

$$f_{00} = f_{11} = f_{22} = f_{33} = 0$$

$$f_{01} = \frac{\partial A_1}{\partial x^0} - \frac{\partial A_0}{\partial x^1} = -\frac{\partial A_x}{\partial t} - \frac{\partial \varphi}{\partial x} = E_x$$

$$f_{02} = \frac{\partial A_2}{\partial x^0} - \frac{\partial A_0}{\partial x^2} = -\frac{\partial A_y}{\partial t} - \frac{\partial \varphi}{\partial y} = E_y$$

$$f_{03} = \frac{\partial A_3}{\partial x^0} - \frac{\partial A_0}{\partial x^3} = -\frac{\partial A_z}{\partial t} - \frac{\partial \varphi}{\partial z} = E_z$$

$$f_{12} = \frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2} = -c \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) = -c(\nabla \times \vec{A})_z = -cB_z$$

$$f_{13} = \frac{\partial A_3}{\partial x^1} - \frac{\partial A_1}{\partial x^3} = -c \left(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z} \right) = c(\nabla \times \vec{A})_y = cB_y$$

$$f_{23} = \frac{\partial A_3}{\partial x^2} - \frac{\partial A_2}{\partial x^3} = -c \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) = -c(\nabla \times \vec{A})_x = -cB_x$$

Zatem:

$$(f_{\mu\nu}) = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix}$$

Do nowego układu inercjalnego przechodzimy dokonując transformacji tensora kowariantnego:

$$f'_{\mu\nu} = \sum_{\mu_1\nu_1} \frac{\partial x^{\mu_1}}{\partial x'^\mu} \frac{\partial x^{\nu_1}}{\partial x'^\nu} f_{\mu_1\nu_1}$$

Podobnie jak w przypadku transformacji tensora metrycznego możemy skorzystać z zapisu macierzowego:

$$(f'_{\mu\nu}) = (\tilde{C}_\mu^\nu)(f_{\mu\nu})(\tilde{C}_\mu^\nu)^T,$$

lub

$$(f'_{\mu\nu}) = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix} \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Po wykonaniu elementarnego mnożenia macierzy uzyskujemy:

$$= \begin{pmatrix} 0 & E_x & \gamma(E_y - \beta cB_z) & \gamma(E_z + \beta cB_y) \\ -E_x & 0 & -\gamma(cB_z - \beta E_y) & \gamma(cB_y + \beta E_z) \\ -\gamma(E_y - \beta cB_z) & \gamma(cB_z - \beta E_y) & 0 & -cB_x \\ -\gamma(E_z + \beta cB_y) & -\gamma(cB_y + \beta E_z) & cB_x & 0 \end{pmatrix}$$

Wynik ten porównujemy z

$$(f'_{\mu\nu}) = \begin{pmatrix} 0 & E'_x & E'_y & E'_z \\ -E'_x & 0 & -cB'_z & cB'_y \\ -E'_y & cB'_z & 0 & -cB'_x \\ -E'_z & -cB'_y & cB'_x & 0 \end{pmatrix}.$$

Możemy więc odczytać:

$$\begin{aligned} E'_x &= E_x & B'_x &= B_x \\ E'_y &= \gamma(E_y - \beta cB_z) & B'_y &= \gamma\left(B_y + \frac{\beta}{c}E_z\right) \\ E'_z &= \gamma(E_z + \beta cB_y) & B'_z &= \gamma\left(B_z - \frac{\beta}{c}E_y\right) \end{aligned}$$

Widzimy więc, że przy transformacji Lorentza współrzędne obu pól elektrycznego i magnetycznego mieszają się nawzajem. Jakościowo wynik ten jest dość oczywisty. Załóżmy, że obserwujemy pole pochodzące od ładunku punktowego. W jego układzie spoczynkowym jest on źródłem wyłącznie pola elektrycznego. W układzie poruszającym się względem niego, ładunek porusza się a taki ładunek jest źródłem również pola magnetycznego.

Podsumowanie VIII.

- Zrobiliśmy pierwszy krok w kierunku elektrodynamiki. Zaproponowaliśmy działanie dla cząstki w zewnętrznym polu elektromagnetycznym.
- Składnik działania opisujący oddziaływanie z polem elektromagnetycznym zaproponowaliśmy w postaci niezmiennika transformacji Lorentza $A_\mu dx^\mu$ zbudowanego z czterowektorowego potencjału pola elektromagnetycznego $A_\mu = (\phi, -c\vec{A})$ i wektora przesunięcia w czasoprzestrzeni $dx^\mu = (cdt, d\vec{x})$.
- Utworzyliśmy funkcję Lagrange'a, otrzymaliśmy wyrażenia na pęd układu (pęd cząstki i pola), energię i funkcję Hamiltona.

- Korzystając z formalizmu Lagrange’a uzyskaliśmy równania ruchu. Żądając odtworzenia wzoru na siłę Lorentza znaleźliśmy związki pomiędzy wektorami pola elektrycznego i magnetycznego oraz czteropotencjałem.

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t}$$

$$\vec{B} = \nabla \times \vec{A}$$

- Przedyskutowaliśmy niezmienniczość cechowania potencjału elektromagnetycznego polegającą na możliwości dodania do czterowektora potencjału czterogradientu dowolnej

$$\text{funkcji określonej w czasoprzestrzeni: } \tilde{A}_\mu = A_\mu - \frac{\partial f}{\partial x^\mu}$$

- Wprowadziliśmy pojęcie tensora pola elektromagnetycznego. Jest nim antysymetryczny tensor, którego 6 niezależnych składowych stanowią składowe pola elektrycznego i magnetycznego w przestrzeni trójwymiarowej.
- Korzystając ze znajomości reguł transformacji tensora znaleźliśmy wyrażenia na zmianę współrzędnych pola elektrycznego i magnetycznego przy przejściach pomiędzy inercjalnymi układami odniesienia.

IX.

Pierwsza para równań Maxwella.

Przypomnijmy uzyskane poprzednio wyrażenia na wektory pola elektrycznego i magnetycznego, które dały nam interpretację składowych czteropotencjału pola elektromagnetycznego:

$$\vec{E} = -\nabla\varphi - \frac{\partial \vec{A}}{\partial t}$$

oraz

$$\vec{B} = \nabla \times \vec{A}$$

Jeżeli nie chcemy posługiwać się potencjałami pola (nie są bezpośrednio mierzalnymi wielkościami fizycznymi) można z uzyskanych równań usunąć potencjały. Podziałajmy na obie strony równań operatorem ∇ :

$$\nabla \times \vec{E} = -\nabla \times \nabla\varphi - \nabla \times \frac{\partial \vec{A}}{\partial t} = -\frac{\partial}{\partial t} \nabla \times \vec{A} = -\dot{\vec{B}}$$

oraz

$$\nabla \cdot \vec{B} = \nabla \cdot \nabla \times \vec{A} = 0$$

Równania te tworzą pierwszą parę równań Maxwella:

Czterowektor gęstości prądu, równanie ciągłości.

Dotychczas uzyskaliśmy aparat matematyczny pozwalający opisać pojedynczą cząstkę w zewnętrznym polu elektromagnetycznym.

Założmy jednak sytuację nieco ogólniejszą umożliwiającą nam konstrukcję teorii nie ograniczającej się do ruchu pojedynczych cząstek. Rozważmy pewien ciągły rozkład ładunków, zadany określoną w przestrzeni funkcją gęstości ładunku $\rho(\vec{r})$ oraz gęstości prądu

$$\vec{j}(\vec{r}) = \rho(\vec{r}) \frac{\partial \vec{r}}{\partial t}$$

i spróbujmy dla nich odtworzyć wyrażenie na działanie.

Wprowadzając gęstość ładunku i prądu musimy zapewnić zachowanie ładunku. Uzyskujemy to żądając spełnienia równania ciągłości:

$$\frac{d}{dt} \int_V \rho dV = - \oint_{\Sigma} \vec{j} d\vec{\sigma}$$

(narysować powierzchnię zamkniętą z zaznaczeniem wektora elementu powierzchni)

Lewa strona równania określa przyrost ładunku w pewnej objętości całkowania V , prawa strumień ładunku wpływający do objętości V przez otaczającą obszar całkowania powierzchnię. Ponieważ zazwyczaj przyjmujemy kierunek wektora elementu powierzchni prostopadły do powierzchni i skierowany na zewnątrz, przed całką umieszczony jest znak minus.

Do prawej strony równania zastosujemy twierdzenie Gaussa słuszne dla dowolnego ciągłego pola wektorowego:

$$\int_V \nabla \cdot \vec{W}(\vec{r}) dV = \oint_{\Sigma} \vec{W}(\vec{r}) d\vec{\sigma}$$

gdzie Σ stanowi zamkniętą powierzchnię otaczającą objętość V .

Przyjmując $\vec{W}(\vec{r}) = \vec{j}(\vec{r})$ otrzymujemy:

$$\frac{d}{dt} \int_V \rho dV = \oint_{\Sigma} \vec{j} d\vec{\sigma} = - \int_V \nabla \cdot \vec{j} dV = - \int_V \text{div} \vec{j} dV,$$

skąd wynika równanie ciągłości w postaci różniczkowej:

$$\int_V \left(\frac{d}{dt} \rho + \text{div} \vec{j} \right) dV = 0 \Rightarrow \text{div} \vec{j} + \frac{\partial \rho}{\partial t} = 0$$

Równanie to można również zapisać w tzw. relatywistycznie niezmienniczej postaci ponieważ gęstość ładunku i gęstość prądu stanowią składowe czterowektora. Można to pokazać prowadząc następujące rozumowanie. Jeżeli zadamy gęstość ładunku ρ , wówczas w elementarnej objętości dV znajduje się ładunek:

$$de = \rho dV.$$

Po pomnożeniu przez elementarne przesunięcie w czasoprzestrzeni otrzymujemy:

$$de dx^\mu = \rho dV dx^\mu = \rho dV dt \frac{dx^\mu}{dt} = \rho \frac{dx^\mu}{dt} d\Omega = j^\mu d\Omega$$

Usuwamy z tego wyrażenia element objętości czasoprzestrzennej:

$$d\Omega = dV dt = dy dz dx dt,$$

który dzięki kompensacji skrócenia odcinka z dylatacją czasu jest skalarem:

$$l' = \frac{1}{\gamma} l_0, t' = \gamma \tau \Rightarrow d\Omega' = d\Omega,$$

uzyskujemy tzw. czterowektor gęstości prądu:

$$j^\mu = \rho \frac{dx^\mu}{dt} = (\rho c, \vec{j}).$$

Jego zerowa składowa jest gęstością ładunku w określonym punkcie przestrzeni pomnożoną przez prędkość światła, a pozostałe trzy składowe wektora $\vec{j} = \rho \vec{v}$ stanowi nierelatywistyczna gęstość prądu. Różniczkowa postać równania ciągłości narzuca na czterowektor gęstości prądu warunek:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = \partial_\mu j^\mu = 0$$

Powróćmy do uzyskanej poprzednio całki działania odpowiedzialnej za oddziaływanie zewnętrznego pola z cząstką relatywistyczną:

$$S_{mf} = -\frac{e}{c} \int_{z_1}^{z_2} A_\mu dx^\mu.$$

Zastąpmy w niej ładunek cząstki e całką z rozkładu gęstości ładunku:

$$e \rightarrow \int_V \rho dV$$

Wówczas całka działania przechodzi w:

$$\begin{aligned} S_{mf} &= -\frac{1}{c} \int_V \rho dV \int_{z_1}^{z_2} A_\mu dx^\mu = -\frac{1}{c} \int_V \rho dV \int_{t_1}^{t_2} A_\mu \frac{dx^\mu}{dt} dt = -\frac{1}{c} \int_\Omega A_\mu \rho \frac{dx^\mu}{dt} dV dt = \\ &= -\frac{1}{c} \int_\Omega A_\mu j^\mu d\Omega \end{aligned}$$

Uzyskaliśmy więc działanie odpowiedzialne za oddziaływanie ciągłego rozkładu gęstości ładunku i prądu danego czterowektorem j^μ z zewnętrznym polem elektromagnetycznym zadany przez czteropotencjał A_μ .

Pole elektromagnetyczne wywołane zadaniem rozkładem ładunków i prądów.

Pole elektryczne i magnetyczne, o którym mówiliśmy było traktowane jako pole zewnętrzne, w którym znajduje się cząstka. Odwróćmy zagadnienie i przeanalizujemy sytuację odwrotną. Potraktujmy cząstkę jako źródło pola. Umożliwi to nam przejście do opisu układów cząstek w polach przez nie wytworzonych. Wtedy pola nie będą musiały być traktowane jako zewnętrzne.

Działanie pola elektromagnetycznego, druga para równań Maxwella.

Musimy utworzyć całkę działania pól swobodnych S_f . Powinno być ono niezmiennikiem transformacji Lorentza kwadratowym w funkcji wektorów pola (równania pola mają być liniowe). Ponieważ czteropotencjał A_μ nie jest jednoznaczna funkcją pola nie nadaje się do tego celu (ze względu na nieliniowość). Można natomiast utworzyć niezmiennik transformacji Lorentza spełniający warunek jednoznaczności z tensora pola elektromagnetycznego. Przyjmijmy:

$$S_f = -\frac{\epsilon_0}{4} \int F_{\mu\nu} F^{\mu\nu} d\Omega$$

ϵ_0 oznacza tzw. przenikalność dielektryczną próżni. Współczynnik został dobrany z wyprzedzeniem tak, by prowadził do poprawnych równań pola.

Całka działania dyskutowanego układu przyjmuje postać:

$$S = S_{mf} + S_f = -\frac{1}{c} \int A_\mu j^\mu d\Omega - \frac{\epsilon_0}{4} \int F_{\mu\nu} F^{\mu\nu} d\Omega = -\frac{1}{c} \int \left(A_\mu j^\mu + \frac{c\epsilon_0}{4} F_{\mu\nu} F^{\mu\nu} \right) d\Omega$$

Żądamy znikania wariacji całki działania

$$\delta S = -\frac{1}{c} \int \left(j^\mu \delta A_\mu + \frac{c\epsilon_0}{4} \delta(F_{\mu\nu} F^{\mu\nu}) \right) d\Omega = -\frac{1}{c} \int \left(j^\mu \delta A_\mu + \frac{c\epsilon_0}{2} F^{\mu\nu} \delta F_{\mu\nu} \right) d\Omega$$

Ponieważ:

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}$$

$$\delta F_{\mu\nu} = \delta \left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} \right) = \frac{\partial \delta A_\nu}{\partial x^\mu} - \frac{\partial \delta A_\mu}{\partial x^\nu}$$

$$\delta S = -\frac{1}{c} \int \left(j^\mu \delta A_\mu + \frac{c\epsilon_0}{2} F^{\mu\nu} \left(\frac{\partial \delta A_\nu}{\partial x^\mu} - \frac{\partial \delta A_\mu}{\partial x^\nu} \right) \right) d\Omega$$

w wyrażeniu tym korzystając z antysymetrii tensora pola elektrycznego i magnetycznego możemy pozbyć się jednego wyrazu:

$$\delta S = -\frac{1}{c} \int \left(j^\mu \delta A_\mu - c\epsilon_0 F^{\mu\nu} \frac{\partial \delta A_\mu}{\partial x^\nu} \right) d\Omega$$

Korzystając z twierdzenia Gaussa w przestrzeni czterowymiarowej:

$$\int_{\Omega} \partial_\nu W^\nu d\Omega = \oint_{\partial\Omega} W^\nu d\sigma_\nu$$

i przyjmując za W :

$$W^\nu = F^{\mu\nu} \delta A_\mu$$

uzyskujemy:

$$\int_{\Omega} \partial_{\nu} (F^{\mu\nu} \delta A_{\mu}) d\Omega = \oint_{\partial\Omega} F^{\mu\nu} \delta A_{\mu} d\sigma_{\nu} = 0$$

Wyraz po prawej stronie znika ze względu na znikanie wariacji na granicy obszaru Ω . Po lewej stronie wykonujemy różniczkowanie pod całką:

$$\int_{\Omega} (F^{\mu\nu} \partial_{\nu} \delta A_{\mu} + \delta A_{\mu} \partial_{\nu} F^{\mu\nu}) d\Omega = 0$$

skąd:

$$\int_{\Omega} F^{\mu\nu} \partial_{\nu} \delta A_{\mu} d\Omega = - \int_{\Omega} \delta A_{\mu} \partial_{\nu} F^{\mu\nu} d\Omega = \int_{\Omega} \delta A_{\mu} \partial_{\nu} F^{\nu\mu} d\Omega$$

Wykorzystując ten związek w wyrażeniu na wariację całki działania otrzymujemy:

$$\delta S = -\frac{1}{c} \int \left(j^{\mu} - c\epsilon_0 \frac{\partial}{\partial x^{\nu}} F^{\nu\mu} \right) \delta A_{\mu} d\Omega$$

Przyrównanie do zera wariacji całki działania dla dowolnej wariacji czteropotencjału daje

$$j^{\mu} - c\epsilon_0 \frac{\partial}{\partial x^{\nu}} F^{\nu\mu} = 0$$

czyli:

$$\frac{\partial}{\partial x^{\nu}} F^{\nu\mu} = \frac{1}{c\epsilon_0} j^{\mu}$$

Otrzymaliśmy 4 równania pola w zapisie tensorowym. Możemy je rozszyfrować wstawiając jawną postać elementów tensora F. W kontrawariantnej wersji:

$$\begin{aligned} (F^{\mu\nu}) &= (g^{\mu\nu})(F_{\mu\nu})(g^{\mu\nu}) = \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \\ &= \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -cB_z & cB_y \\ E_y & cB_z & 0 & -cB_x \\ E_z & -cB_y & cB_x & 0 \end{pmatrix} \end{aligned}$$

$$\frac{\partial}{\partial x^{\mu}} F^{\mu 0} = \frac{1}{c\epsilon_0} j^0 \Leftrightarrow \nabla \vec{E} = \frac{1}{\epsilon_0} \rho$$

$$\frac{\partial}{\partial x^{\mu}} F^{\mu 1} = \frac{1}{c\epsilon_0} j^1 \Leftrightarrow -\frac{1}{c} \frac{\partial E_x}{\partial t} + c \frac{\partial B_z}{\partial y} - c \frac{\partial B_y}{\partial z} = \frac{1}{c\epsilon_0} j^x$$

$$\frac{\partial}{\partial x^{\mu}} F^{\mu 2} = \frac{1}{c\epsilon_0} j^2 \Leftrightarrow -\frac{1}{c} \frac{\partial E_y}{\partial t} - c \frac{\partial B_z}{\partial x} + c \frac{\partial B_x}{\partial z} = \frac{1}{c\epsilon_0} j^y$$

$$\frac{\partial}{\partial x^{\mu}} F^{\mu 3} = \frac{1}{c\epsilon_0} j^3 \Leftrightarrow -\frac{1}{c} \frac{\partial E_z}{\partial t} + c \frac{\partial B_y}{\partial x} - c \frac{\partial B_x}{\partial y} = \frac{1}{c\epsilon_0} j^z$$

Co można przepisać jako:

$\epsilon_0 \nabla \vec{E} = \rho$ oraz $-\epsilon_0 \dot{\vec{E}} + c^2 \epsilon_0 \nabla \times \vec{B} = \vec{j}$
 a po wykorzystaniu znanych związków:

$$c^2 = \frac{1}{\epsilon_0 \mu_0}, \epsilon_0 \vec{E} = \vec{D}, \mu_0 \vec{H} = \vec{B}$$

dostajemy:

$$\nabla \vec{D} = \rho$$

$$\nabla \times \vec{H} = \vec{D} + \vec{j}$$

Są to kolejne dwa równania Maxwella.

Równania Maxwella w postaci różniczkowej.

Uzyskaliśmy właśnie dwa równania dające związki pomiędzy polami magnetycznym i elektrycznym a ich źródłami gęstością ładunku i prądu. Wraz z równaniami uzyskanymi poprzednio tworzą one komplet równań Maxwella stanowiących bazę dla całej elektrodynamiki. Zestawmy je razem:

1. $\nabla \times \vec{E} = -\dot{\vec{B}}$
2. $\nabla \cdot \vec{B} = 0$
3. $\nabla \cdot \vec{D} = \rho$
4. $\nabla \times \vec{H} = \vec{D} + \vec{j}$

Do równań Maxwella dorzucić należy równanie ciągłości, które pojawiło się po wprowadzeniu gęstości ładunku i prądu:

$$5. \operatorname{div} \vec{j} + \frac{\partial \rho}{\partial t} = 0$$

Kowariantny (jawnie relatywistyczny) zapis równań Maxwella.

Równania Maxwella można również zapisać w jawnie relatywistycznej postaci. Kowariantny zapis dwóch ostatnich równań Maxwella już znamy:

$$\partial_\nu F^{\nu\mu} = \frac{1}{c\epsilon_0} j^\mu$$

Dwa pierwsze równania można zapisać podobnie:

$$\partial_\nu G^{\nu\mu} = 0,$$

gdzie G nosi nazwę tensora dualnego, a jego składowe wyrażają się podobnie jak w przypadku tensora F przez składowe pola elektrycznego i magnetycznego:

$$(G^{\mu\nu}) = \begin{pmatrix} 0 & cB_x & cB_y & cB_z \\ -cB_x & 0 & -E_z & E_y \\ -cB_y & E_z & 0 & -E_x \\ -cB_z & -E_y & E_x & 0 \end{pmatrix}$$

Po podstawieniu do poprzedniego wzoru otrzymujemy:

$$\frac{\partial}{\partial x^\mu} G^{\mu 0} = 0 \Leftrightarrow \nabla \vec{B} = 0$$

oraz

$$\frac{\partial}{\partial x^\mu} G^{\mu 1} = 0 \Leftrightarrow \frac{\partial B_x}{\partial t} + \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = 0$$

$$\frac{\partial}{\partial x^\mu} G^{\mu 2} = 0 \Leftrightarrow \frac{\partial B_y}{\partial t} - \frac{\partial E_z}{\partial x} + \frac{\partial E_x}{\partial z} = 0,$$

$$\frac{\partial}{\partial x^\mu} G^{\mu 3} = 0 \Leftrightarrow \frac{\partial B_z}{\partial t} + \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = 0$$

co w zapisie wektorowym daje

$$\dot{\vec{B}} = -\nabla \times \vec{E}.$$

Również znamy odpowiedni zapis równania ciągłości:

$$\partial_\mu j^\mu = 0.$$

Podsumowanie IX.

- Uzyskaliśmy pierwszą parę równań Maxwella:

$$\nabla \times \vec{E} = -\dot{\vec{B}}$$

$$\nabla \cdot \vec{B} = 0$$

- Wprowadziliśmy pojęcie czterowektora gęstości prądu: $j^\mu = (\rho c, \vec{j})$
- Narzuciliśmy na ten czterowektor warunek ciągłości: $\partial_\mu j^\mu = 0$
- 5. Uzyskaliśmy wyrażenie na działanie wynikające z oddziaływania gęstości ładunku i prądu z zewnętrznym polem elektromagnetycznym:

$$S_{mf} = -\frac{1}{c} \int_{\Omega} A_\mu j^\mu d\Omega$$

- Przy pomocy tensora pola elektrycznego i magnetycznego utworzyliśmy niezmiennik, pozwalający na uzyskanie z zasady najmniejszego działania opisu pól, których źródłami są (traktowane jako zewnętrzne) gęstości ładunków i gęstości prądów:

$$S_f = -\frac{\epsilon_0}{4} \int F_{\mu\nu} F^{\mu\nu} d\Omega$$

- Uzyskaliśmy drugą parę równań Maxwella.

$$\nabla \cdot \vec{D} = \rho$$

$$\nabla \times \vec{H} = \dot{\vec{D}} + \vec{j}$$

- Zebraliśmy razem równania Maxwella i równanie ciągłości w różniczkowej i relatywistycznie niezmienniczej (kowariantnej) postaci.

X.

Równania Maxwella w postaci całkowej.

W wielu zastosowaniach znacznie wygodniej jest posługiwać się równaniami Maxwella w postaci całkowej, którą uzyskuje się stosując odpowiednio cytowane już wcześniej twierdzenie Gaussa pozwalające na zamianę całki po objętości na całkę po powierzchni tę objętość otaczającej:

$$\int_V \nabla \vec{W} dV = \int_{\Sigma} \vec{W} \cdot d\vec{\sigma},$$

lub twierdzenie Stockesa, które również jest słuszne dla dowolnej funkcji wektorowej:

$$\oint_C \vec{W} \cdot d\vec{l} = \iint_S \nabla \times \vec{W} \cdot d\vec{S}$$

i pozwala zamieniać całki po powierzchni S na całki po jej obwodzie oznaczonym jako kontur.

Pierwsze równanie Maxwella

$$1. \quad \nabla \times \vec{E} = -\dot{\vec{B}}$$

przekształcamy przyjmując w równaniu Stockesa $\vec{W}(\vec{r}) = \vec{E}(\vec{r})$:

$$\oint_C \vec{E} \cdot d\vec{l} = \iint_S \nabla \times \vec{E} \cdot d\vec{S} = -\iint_S \dot{\vec{B}} \cdot d\vec{S} = \frac{d}{dt} \iint_S \vec{B} \cdot d\vec{S}.$$

Drugie równanie:

$$2. \quad \nabla \cdot \vec{B} = 0$$

przekształcamy podstawiając $\vec{W}(\vec{r}) = \vec{B}(\vec{r})$ w równaniu Gaussa

$$\int_{\Sigma} \vec{B} \cdot d\vec{\sigma} = \int_V \nabla \cdot \vec{B} dV = 0.$$

Trzecie równanie

$$3. \quad \nabla \cdot \vec{D} = \rho$$

otrzymujemy z równania Gaussa podstawiając $\vec{W}(\vec{r}) = \vec{D}(\vec{r})$:

$$\int_{\Sigma} \vec{D} \cdot d\vec{\sigma} = \int_V \nabla \cdot \vec{D} dV = \int_V \rho dV = q,$$

gdzie q stanowi ładunek objęty powierzchnią całkowania Σ .

Czwarte równanie:

$$4. \quad \nabla \times \vec{H} = \dot{\vec{D}} + \vec{j}$$

przekształcamy znowu przy pomocy równania Stockesa podstawiając $\vec{W}(\vec{r}) = \vec{H}(\vec{r})$:

$$\begin{aligned} \oint_C \vec{H} \cdot d\vec{l} &= \iint_S \nabla \times \vec{H} \cdot d\vec{S} = \iint_S (\dot{\vec{D}} + \vec{j}) \cdot d\vec{S} = \\ &= \frac{d}{dt} \iint_S \vec{D} \cdot d\vec{S} + \iint_S \vec{j} \cdot d\vec{S} = \frac{d}{dt} \iint_S \vec{D} \cdot d\vec{S} + I \end{aligned}$$

gdzie przez I oznaczamy prąd ograniczony konturem C.

Zestawione razem całkowe odpowiedniki różniczkowych równań Maxwella i równania ciągłości mają postać:

$$1. \quad \oint_C \vec{E} \cdot d\vec{l} = \frac{d}{dt} \iint_S \vec{B} \cdot d\vec{S}$$

$$2. \quad \oint_{\Sigma} \vec{B} \cdot d\vec{\sigma} = 0$$

$$3. \oint_{\Sigma} \vec{D} \cdot d\vec{\sigma} = q$$

$$4. \oint_C \vec{H} \cdot d\vec{l} = \frac{d}{dt} \iint_S \vec{D} d\vec{S} + I$$

$$5. \frac{d}{dt} \int_V \rho dV = - \oint_{\partial V} \vec{j} \cdot d\vec{\sigma}$$

Refleksja nad sposobem wprowadzenia równań Maxwella.

Równania Maxwella wraz z równaniem ciągłości pozwalają opisać wszystkie zjawiska elektrodynamiki. Wystarczy je tylko rozwiązać. Zauważmy, że udało nam się je wyprowadzić wykorzystując zasadę najmniejszego działania startując z mechaniki relatywistycznej. Po drodze zainwestowaliśmy naprawdę niewielką informację fizyczną. Właściwie tylko taką, która prowadziła do ujednolicenia tworzonej teorii ze znaną wcześniej elektrodynamiką. Maxwell jednakże swoje równania zaproponował nie znając mechaniki relatywistycznej. Jego teoria powstała na bazie kilku sformułowanych wcześniej praw eksperymentalnych i były ich syntezą oraz uogólnieniem. Jako ciekawostkę można podać, że „czwarte” prawo Maxwella, które wyprowadzone zostało z prawa Ampera, w oryginale nie zawierało pochodnej pola elektrycznego

$$\nabla \times \vec{H} = \vec{j},$$

ponieważ eksperyment, który pozwolił na jego sformułowanie był prowadzony przy stałych prądach. Maxwell formułując swoje równania zauważył, asymetrię tego równania z równaniem „pierwszym”.

$$\nabla \times \vec{E} = -\dot{\vec{B}}$$

Ze względu na występujący w obu równaniach iloczyn wektorowy, gradient lewych stron obu równań musi się zerować. Policzmy gradient prawych stron obu równań. Dla równania „pierwszego”:

$$\nabla \cdot \nabla \times \vec{E} = -\nabla \cdot \dot{\vec{B}} = 0$$

zeruje się dzięki równaniu „drugiemu”. W przypadku Prawa Ampera w oryginalnym brzmieniu:

$$\nabla \cdot \nabla \times \vec{H} = \nabla \cdot \vec{j}$$

dywergencja gęstości prądu na ogół nie zeruje się. Korzystając z całkowicie teoretycznych rozważań (symetria równań) Maxwell uogólnił równanie Ampera dodając do jego prawej strony czasową pochodną pola elektrycznego. Uzyskał dzięki temu zerowanie się prawej strony równania przy wykorzystaniu trzeciego prawa Maxwella i równania ciągłości.

$$\nabla \cdot \nabla \times \vec{H} = \nabla \cdot \vec{j} + \nabla \cdot \dot{\vec{D}} = \nabla \cdot \vec{j} + \frac{\partial}{\partial t} \rho = 0$$

Poprawka ta potwierdziła się w eksperymentach wykraczających poza megnetostatykę. W historycznej kolejności najpierw pojawiły się prawa elektrodynamiki sformułowane na bazie eksperymentu, następnie prawa teoretyczne Maxwella a na końcu fizyka relatywistyczna. My wyszliśmy od relatywistyki, uzyskaliśmy z niej prawa Maxwella, żeby uzupełnić brakujące ogniwo, uzyskamy z praw Maxwella eksperymentalne prawa elektrodynamiki.

Zastosowanie równań Maxwella w postaci całkowej.

Rozwiązywanie układów równań różniczkowych cząstkowych jakimi są równania Maxwella jest na ogół dosyć żmudne i kłopotliwe. W wielu przypadkach a w szczególności do opisu

układów o wysokiej symetrii znacznie łatwiej jest zastosować ich całkową postać. Rozważmy kilka takich przykładów.

Prawo Coulomba.

Jako pierwszy przykład zastosowania równań Maxwella znajdziemy rozkład pola wokół punkowego ładunku. Umieścimy w początku układu ładunek q . Ponieważ wyróżnia on tylko jeden punkt w przestrzeni rozkład pól powinien mieć symetrię sferyczną. Oczekujemy pola $\vec{D}$ skierowanego zgodnie z promieniem wodzącym punktu w którym mierzymy pole

$$\vec{D} = D \frac{\vec{r}}{r}$$

Wybieramy powierzchnię sferyczną o promieniu r i korzystamy z równania

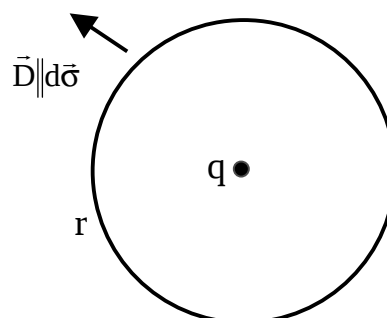
$$\oint_{\Sigma} \vec{D} \cdot d\vec{\sigma} = q$$

Wektor $d\vec{\sigma}$ jest prostopadły do powierzchni sfery Σ i skierowany na zewnątrz, jest więc równoległy do wektora $\vec{D}$. Całka powierzchniowa po lewej stronie równania jest łatwa do wyliczenia i przyjmuje wartość $4\pi r^2 D$. Po prawej stronie równania występuje całkowity ładunek zawarty w objętości objętej powierzchnią Σ .

$$4\pi r^2 D = q \Rightarrow D = \frac{q}{4\pi r^2}$$

a w zapisie wektorowym:

$$\vec{D}(r) = \frac{q}{4\pi r^2} \frac{\vec{r}}{r} \text{ lub } \vec{E}(r) = \frac{q}{4\pi\epsilon_0 r^2} \frac{\vec{r}}{r}$$



Potencjał skalarny pochodzący od ładunku punkowego.

Znając rozkład pola elektrostatycznego możemy wyliczyć jego potencjał. Pamiętamy związek pomiędzy wektorem natężenia pola elektrostatycznego a czteropotencjałem:

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t}$$

Jeżeli założymy brak czasowej zależności pól, pozostaje

$$\vec{E} = -\nabla\phi$$

Możemy więc wyliczyć potencjał skalarny pola we wszystkich przypadkach, w których wyliczyliśmy rozkład pola elektrycznego. Zauważmy, że przez założenie niezależności czasowej potencjału wektorowego, z niejednoznaczności czteropotencjału pozostała tylko możliwość dodania dowolnej stałej, czyli przyjęcia poziomu odniesienia dla potencjału skalarnego.

$$\vec{\tilde{A}} = \vec{A} + \frac{1}{c} \nabla f$$

$$\tilde{\phi} = \phi - \frac{1}{c} \frac{\partial f}{\partial t}$$

Jeżeli pole magnetyczne ma nie zależeć od czasu, to

$$\frac{\partial}{\partial t} \vec{\tilde{A}} = \frac{\partial}{\partial t} \vec{A} + \frac{1}{c} \frac{\partial}{\partial t} \nabla f = \frac{1}{c} \frac{\partial}{\partial t} \nabla f = 0$$

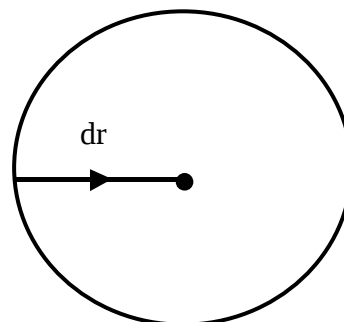
Zatem funkcja f nie zależy od czasu albo zmiennych przestrzennych. W pierwszym przypadku potencjał skalarny nie ulega zmianie, w przypadku drugim do potencjału skalarnego dodajemy stałą.

Zazwyczaj jako zerowy przyjmujemy potencjał w nieskończoności. Policzmy rozkład potencjału pochodzącego od ładunku punktowego. Ze względu na symetrię sferyczną scałkujemy potencjał poruszając się wzdłuż promienia od nieskończoności (gdzie potencjał się zeruje) do punktu, w którym potencjał wyznaczamy.

$$d\phi = d\vec{r} \cdot \nabla\phi = -d\vec{r} \cdot \vec{E} = -E dr$$

$$\phi(r) = \int d\phi = -\int_{\infty}^r E(r') dr' = -\int_{\infty}^r \frac{q}{4\pi\epsilon_0 r'^2} dr' = \left. \frac{q}{4\pi\epsilon_0 r'} \right|_{\infty}^r = \frac{q}{4\pi\epsilon_0 r}$$

Uzyskaliśmy zatem prawo, które Coulomb wyznaczył eksperymentalnie.



Układ kilku ładunków punktowych.

W przypadku kilku ładunków punktowych będących źródłami pola układ nie posiada wysokiej symetrii i stosowanie w takim przypadku prawa Gaussa nie ułatwia rozwiązania. Stosujemy wtedy zasadę superpozycji pól. Liniowość równań Maxwella względem źródeł pola pozwala liczyć oddzielnie rozkłady pól pochodzących od różnych źródeł i dodawanie ich. Zarówno wektory pola jak i potencjały są wielkościami addytywnymi.

$$\phi(\vec{r}) = \sum_i \frac{q_i}{4\pi\epsilon_0 |\vec{r} - \vec{r}_i|} \quad \vec{E}(\vec{r}) = \sum_i \frac{q_i (\vec{r} - \vec{r}_i)}{4\pi\epsilon_0 |\vec{r} - \vec{r}_i|^3}$$

gdzie $\vec{r}$ jest wektorem wodzącym punktu, w którym liczymy pole, a $\vec{r}_i$ położeniem ładunku q_i .

Ciągły rozkład gęstości ładunku.

W przypadku gdy mamy zadany ciągły rozkład ładunku w przestrzeni, oraz zażądamy znikania potencjału w nieskończoności możemy skorzystać z wyrażenia na potencjał pochodzący od układu ładunków punktowych dokonując przejścia:

$$q_i \rightarrow \rho(\vec{r}') dV = \rho(\vec{r}') d^3r'$$

Sumę po położeniach ładunków zastępujemy całką po rozkładzie gęstości i rozwiązanie przyjmuje postać:

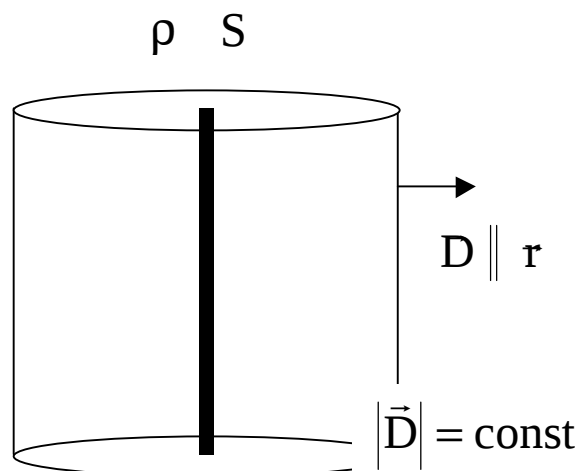
$$\phi(\vec{r}) = \int \frac{\rho(\vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|} d^3r' \quad \vec{E}(\vec{r}) = \int \frac{\rho(\vec{r}') (\vec{r} - \vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|^3} d^3r'$$

Pole elektryczne wokół jednorodnie naładowanego walca.

Przedyskutujmy teraz podobny problem, ale dla układu posiadającego symetrię cylindryczną.

Rozważmy nieskończenie długi jednorodnie naładowany gęstością przestrzenną ładunku ρ pręt, o powierzchni przekroju S . Wyliczymy rozkład pola elektrycznego na zewnątrz pręta w odległości r od jego osi. W tym celu otaczamy pręt powierzchnią cylindryczną o promieniu podstawy r i zadanej długości l . Pole elektryczne jest prostopadłe do pobocznic cylindra, i równoległe do jego podstaw.

Wkład do całki powierzchniowej daje pole jedynie na pobocznicy



$$2\pi r l D = \rho S l$$

zatem

$$\vec{D} = \frac{\rho S}{2\pi r} \frac{\vec{r}}{r}$$

gdzie promień wodzący $\vec{r}$ jest prostopadły do osi walca.

Symetria prostokątna- jednorodnie naładowana płaszczyzna, kondensator.

Jednorodne pole pochodzące od płaszczyzny naładowanej jednorodnym ładunkiem powierzchniowym σ . Wybieramy powierzchnię prostopadłościenną obejmującą część płaszczyzny i zawarty na niej ładunek. Oczekujemy wektora natężenia pola w kierunku prostopadłym do naładowanej powierzchni. Całkowanie po czterech ściankach prostopadłościanu równoległych do pola daje zero a po dwóch ściankach prostopadłych do pola daje

$$2SD = q$$

Ładunek q z kolei wyrażony przez gęstość powierzchniową jest:

$$q = S\sigma$$

Ostatecznie:

$$D = \frac{1}{2} \sigma.$$

Na naładowanej powierzchni wektor D doznaje skoku o ładunku powierzchniowego.

Pole wytworzone przez nieskończoną płaszczyznę jest jednorodne i nie zależy od odległości. Zupełnie podobnie rozważamy problem kondensatora płaskiego. Wstawmy drugą płaszczyznę naładowaną ładunkiem

powierzchniowym $-\sigma$ (dla zachowania neutralności ładunkowej). Co się zmieniło. Jeżeli zbudujemy prostopadłościan obejmujący obie powierzchnie ładunek w nim zawarty się zeruje. Na zewnątrz kondensatora pole jest zerowe. Prostopadłościan obejmujący tylko jedną powierzchnię obejmuje ładunek:

$$q = S\sigma$$

Pole elektrostatyczne pomiędzy okładkami

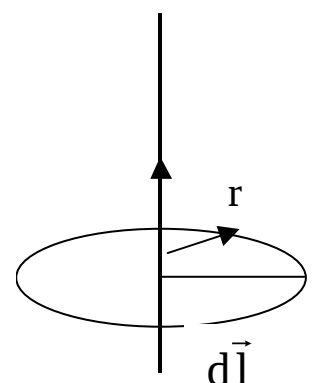
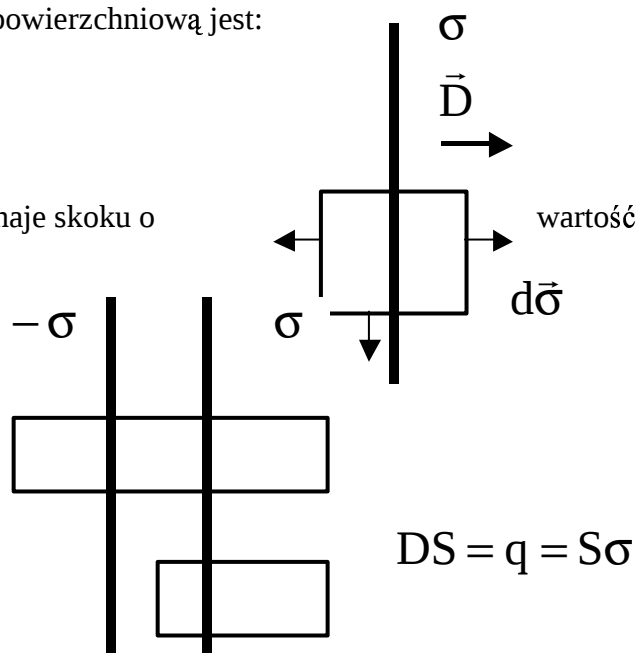
$$D = \sigma$$

Rozwiązanie dla kondensatora możemy również uzyskać z zasady superpozycji.

Oczywiście, żeby przybliżenie nieskończenie dużych płaszczyzn (okładek) było poprawne, odległość pomiędzy okładkami musi być dużo mniejsza od ich zewnętrznych rozmiarów, żeby mogły być zaniedbane efekty niejednorodności pola na brzegach kondensatora.

Zastosowania praw Maxwella w postaci całkowej do magnetostatyki.

Przedyskutowany poprzednio układ o symetrii cylindrycznej wykorzystajmy do rozwiązania nieco odmiennego problemu. Załóżmy, że przez pręt płynie prąd elektryczny o gęstości $\vec{j}$, czyli natężeniu $I = jS$. Gęstości prądu lub zmienne pola elektryczne są źródłami pola magnetycznego. Odpowiednie równanie Maxwella



$$\oint_C \vec{H} \cdot d\vec{l} = \frac{d}{dt} \iint_S \vec{D} d\vec{\sigma} + I$$

ograniczone do magnetostatyki (zakładamy brak zależności czasowej pól) ma postać:

$$\oint_C \vec{H} \cdot d\vec{l} = I$$

Możemy je wykorzystać do policzenia natężenia pola magnetycznego wokół nieskończenie długiego prostoliniowego przewodnika w którym płynie prąd o zadanym natężeniu I . Ponieważ układ posiada symetrię cylindryczną, natężenie indukowanego pola magnetycznego zależy wyłącznie od odległości r od przewodnika. Otaczamy więc przewodnik okręgiem o promieniu r położonym w płaszczyźnie prostopadłej do przewodnika. Wówczas $d\vec{l}$ jest równoległe do $\vec{H}$ i równanie Maxwella daje:

$$2\pi r H = I$$

Zatem

$$H = \frac{I}{2\pi r}$$

Pozostaje ustalenie zwrotu wektora $\vec{H}$. Musimy zapamiętać, że kierunek całkowania po konturze $d\vec{l}$ wybieramy prawoskrętnie względem elementu całkowania po powierzchni $d\vec{\sigma}$. Jeżeli więc płaszczyznę ($d\vec{\sigma}$) skierujemy zgodnie z kierunkiem prądu, to wektor $\vec{H}$ będzie miał zwrot zgodny z kierunkiem śruby prawoskrętnej.

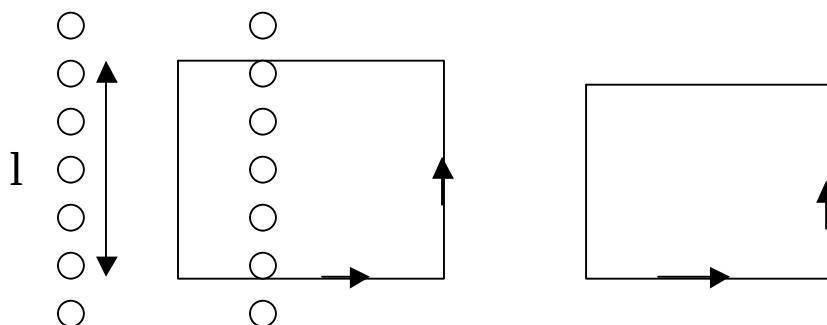
Pole magnetyczne wewnątrz nieskończenie długiego solenoidu.

Wewnątrz i na zewnątrz nieskończenie długiego solenoidu pole magnetyczne powinno być równoległe do jego osi. Korzystamy z całkowitej postaci równania

$$\oint_C \vec{H} \cdot d\vec{l} = I$$

$$I = l n i$$

Ponieważ żaden kontur całkowania przeprowadzony na zewnątrz solenoidu nie obejmuje prądu, pole magnetyczne musi być równe zero. Dla wyznaczenia pola wewnątrz solenoidu wybieramy kontur



całkowania przebiegający równoległe do jego osi wybranej długości l , następnie prostopadle przecinający jego brzozy i zamykamy go dowolną linią na zewnątrz solenoidu. Kontur obejmuje łącznie prąd $l \cdot n \cdot i$, gdzie n oznacza liczbę zwojów przypadających na jednostkę długości solenoidu. Całka po konturze

$$\oint_C \vec{H} \cdot d\vec{l} = H \cdot l = l \cdot n \cdot i$$

zatem

$$H = n \cdot i$$

i nie zależy od odległości od osi solenoidu. Pole magnetyczne w jego wnętrzu jest jednorodne.

Przedstawione przykłady zastosowań równań Maxwella w postaci całkowitej nie wyczerpują wszystkich możliwości, ale ilustrują sposób ich stosowania.

Podsumowanie X.

- Korzystając z praw Gaussa i Stockesa znaleźliśmy całkowite odpowiedniki różniczkowych równań Maxwella.
- Postać całkowa jest dużym udogodnieniem, jeżeli układ ma wysoką symetrię.
- Korzystając z całkowej postaci równań Maxwella wyprowadziliśmy prawo Coulomba - wyliczyliśmy rozkład potencjału wokół ładunku punkowego. Był to problem o symetrii sferycznej.
- Wynik dzięki zasadzie superpozycji uogólniliśmy na układ ładunków punktowych a następnie na ciągły rozkład ładunku.
- Dla symetrii cylindrycznej wyliczyliśmy rozkład pola elektrycznego wokół jednorodnie naładowanego nieskończenie długiego pręta oraz rozkład pola magnetycznego wokół prostoliniowego przewodnika z prądem, znany jako prawo Ampera.
- Dla symetrii prostokątnej wyliczyliśmy pole elektryczne wokół naładowanej powierzchni i w kondensatorze płaskim, oraz pole magnetyczne wewnątrz nieskończenie długiego solenoidu.

XI.

Transformacja Lorentza w magnetostatyce.

Dyskusję zastosowań równań Maxwella w układach o wysokiej symetrii zakończę dygresją z zakresu teorii relatywistycznej.

Omówiony wcześniej problem pola magnetycznego pochodzącego od przewodnika z prądem możemy rozwiązać również inną ciekawą metodą, korzystając z transformacji Lorentza dla pól elektrycznego i magnetycznego. Załóżmy, że prąd I jest spowodowany przepływem jednorodnego ładunku o gęstości ρ poruszającego się z jednostajną prędkością V . Możemy wyliczyć natężenie pola elektrycznego pochodzącego od gęstości ρ w układzie odniesienia w którym ładunek spoczywa, czyli poruszającego się z prędkością V . Następnie korzystając z transformacji Lorentza możemy przeprowadzić pola do układu spoczynkowego. Skierujmy układ odniesienia osią x w kierunku prędkości prądu, oś y w kierunku promienia wodzącego (łączącego przewodnik z punktem w którym liczymy pole) $\vec{r}$. Korzystając z rozwiązane go wcześniej problemu elektrostatyki otrzymujemy niezerową jedynie składową

$$D_y = \frac{\rho S}{2\pi r}$$

Pozostałe składowe pola elektrycznego, oraz pole magnetyczne są zerowe.

Przejście do układu spoczywającego dokonujemy przy pomocy uzyskanych wcześniej wzorów. Uzyskujemy niezerowe

$$E'_y = \gamma E_y - \gamma \beta c B_z = \gamma E_y \quad i$$

$$B'_z = \gamma B_z - \gamma \frac{\beta}{c} E_y = \gamma \frac{\beta}{c} E_y$$

Jeżeli problem nie jest relatywistyczny $\gamma = 1$

$$E'_y = E_y$$

i

$$B'_z = \frac{\beta}{c} E_y = \frac{V}{c^2} E_y = \mu_0 \epsilon_0 V E_y = \mu_0 V D_y = \mu_0 V \frac{\rho S}{2\pi r} = \mu_0 \frac{j S}{2\pi r} = \frac{\mu_0 I}{2\pi r} \text{ lub}$$

$$H_z = \frac{I}{2\pi r}$$

Jeżeli przewodnik jest ładunkowo neutralny, poruszającemu się ładunkowi musi towarzyszyć spoczywający ładunek przeciwny. Korzystając z zasady superpozycji, pole pochodzące od niego liczymy w jego układzie spoczynkowym, czyli docelowym. Jest on źródłem pola elektrostatycznego, które neutralizuje składową E'_y .

Wynik jest identyczny z uzyskanym z drugiego prawa Maxwella. Świadczy to o wewnętrznej spójności omawianej teorii.

Zastosowania równań Maxwella w postaci różniczkowej.

Nie zawsze układ źródeł pola elektromagnetycznego jest na tyle symetryczny, by całkowita postać praw Maxwella mogła być wykorzystana. Musimy wtedy odwołać się do ich różniczkowej postaci.

Równanie Poissona i Laplace'a.

Powróćmy do problemu statycznego pola elektrycznego pochodzącego od ciągłego rozkładu ładunku. Poszukajmy metody pozwalającej na wyliczenie rozkładu pola w układach, w których symetria jest zbyt niska by mogły być wykorzystane całkowite równania Maxwella. Wykorzystamy ich postać różniczkową. Przypomnijmy sobie związki:

$$\nabla \cdot \vec{D} = \rho(\vec{r})$$

oraz

$$\vec{E} = -\nabla\varphi(\vec{r})$$

Łącząc je otrzymujemy równanie różniczkowe drugiego stopnia na potencjał skalarny:

$$\nabla^2\varphi(\vec{r}) = -\frac{1}{\epsilon_0}\rho(\vec{r})$$

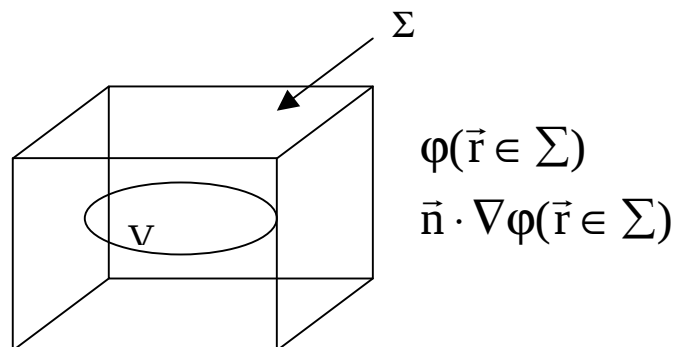
Jest to równanie **Poissona** pojawiające się również w innych działach fizyki. W obszarach przestrzeni pozbawionych ładunku równanie to przechodzi w równanie Laplace'a:

$$\nabla^2\varphi(\vec{r}) = 0.$$

Jesteśmy zmuszeni szukać rozwiązań równania Poissona lub Laplace'a we wszystkich przypadkach w których nie uda się nam skorzystać z całkowitej postaci odpowiedniego równania Maxwella.

Ponieważ równania te są równaniami różniczkowymi drugiego rzędu na pochodne cząstkowe, ich rozwiązanie wymaga zainwestowania warunków brzegowych, które mogą być zadane w postaci wartości pola -warunki brzegowe Dirichleta, lub jego pochodnej normalnej -warunki brzegowe Neumanna na powierzchni zamkniętej otaczającej obszar przestrzeni, w którym liczymy rozkład pola.

W przypadku numerycznego rozwiązywania równania Poissona lub Laplace'a prawidłowe wyznaczenie warunków brzegowych jest kluczem do rozwiązania postawionego problemu.



Zauważmy, że znamy rozkład potencjału pochodzącego od ciągłego rozkładu ładunku przy założeniu zerowania się potencjału w nieskończoności:

$$\varphi(\vec{r}) = \int \frac{\rho(\vec{r}')}{4\pi\epsilon_0|\vec{r} - \vec{r}'|} d^3r'$$

Wynik ten uzyskaliśmy uogólniając prawo Coulomba. Sprawdźmy, czy uzyskany wynik spełnia równanie Poissona. Wykorzystamy przy tym tzw. deltę Diraca. Jest to funkcja, której własności definiujemy poprzez całkę:

$$\delta(\vec{r} - \vec{r}') = \int d^3k e^{i\vec{k}(\vec{r} - \vec{r}')} \quad \text{oraz}$$

oraz

$$\int d^3r \delta(\vec{r}) = 1$$

druga własność wynika z pierwszej. Często używane przedstawienie całkowite funkcji delta Diraca ma postać:

$$\delta(\vec{r} - \vec{r}') = \int d^3k e^{i\vec{k}(\vec{r} - \vec{r}')} \quad \text{oraz}$$

Napiszmy również potencjał kulombowski w postaci transformaty Fouriera:

$$\frac{1}{r} = \int d^3k \frac{4\pi}{k^2} e^{i\vec{k}(\vec{r} - \vec{r}')} \quad \text{oraz}$$

Wstawmy zatem wyrażenie na potencjał do równania Poissona:

$$\begin{aligned}\nabla^2 \phi(\vec{r}) &= \nabla_r^2 \int \frac{\rho(\vec{r}')}{4\pi\epsilon_0 |\vec{r} - \vec{r}'|} d^3r' = \frac{1}{4\pi\epsilon_0} \nabla_r^2 \int d^3r' \rho(\vec{r}') \int d^3k \frac{4\pi}{k^2} e^{i\vec{k}(\vec{r} - \vec{r}')} = \\ &= \frac{1}{\epsilon_0} \int d^3r' \rho(\vec{r}') \int d^3k \frac{1}{k^2} \nabla^2 e^{i\vec{k}(\vec{r} - \vec{r}')} = -\frac{1}{\epsilon_0} \int d^3r' \rho(\vec{r}') \int d^3k e^{i\vec{k}(\vec{r} - \vec{r}')} = \\ &= -\frac{1}{\epsilon_0} \int d^3r' \rho(\vec{r}') \delta(\vec{r} - \vec{r}') = -\frac{1}{\epsilon_0} \rho(\vec{r})\end{aligned}$$

Twierdzenie o jednoznaczności rozwiązań równania Laplace'a i Poissona.

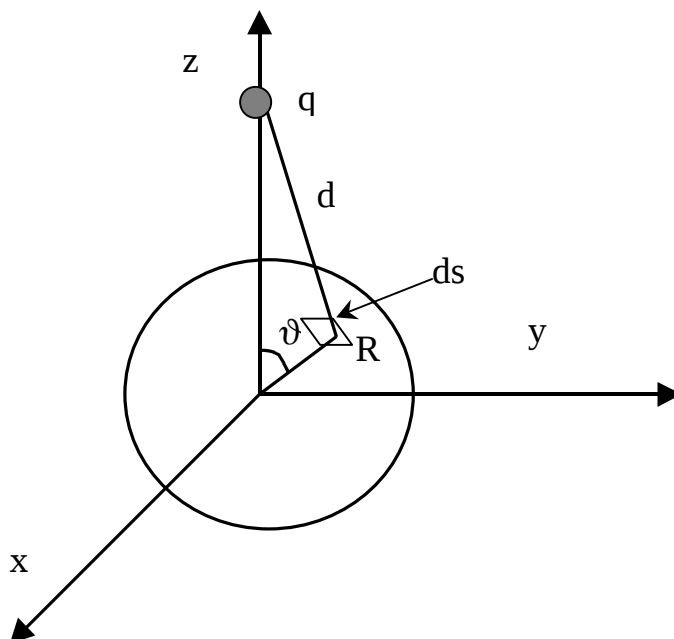
Dość ważną dla późniejszych zastosowań własność równania Laplace'a można sformułować jako twierdzenie o jednoznaczności:

1. Rozwiązanie równania Laplace'a w pewnej objętości V jest określone jednoznacznie jeżeli spełnia warunek brzegowy na powierzchni zamkniętej otaczającej tę objętość.

W celu udowodnienia tego twierdzenia musimy najpierw wykazać inną własność równania Laplace'a:

2. Wartość potencjału elektrostatycznego w wybranym punkcie $\vec{r}$ przestrzeni jest średnią potencjału liczoną po sferze o środku w punkcie $\vec{r}$ i dowolnym promieniu.

Aby to pokazać zbadamy potencjał elektrostatyczny w początku układu wytworzony przez ładunek punktowy q umieszczony w punkcie $(0,0,z)$. Otaczamy początek układu sferą o środku położonym w początku układu i promieniu $R < z$. Odległość ładunku od punktu na sferze, którego wektor wodzący tworzy z osią „ z ” kąt ϑ oznaczmy przez d .



W tym punkcie potencjał wytworzony przez ładunek dany jest wyrażeniem:

$$\phi = \frac{1}{4\pi\epsilon_0} \frac{1}{d}$$

Wyliczmy wartość średnią potencjału całkując ją po powierzchni sfery:

$$\bar{\phi} = \frac{1}{4\pi R^2} \frac{q}{4\pi\epsilon_0} \oint \frac{1}{d} ds = \frac{1}{4\pi R^2} \frac{q}{4\pi\epsilon_0} \int_0^{2\pi} d\phi \int_0^\pi d\vartheta \frac{R^2 \sin \vartheta}{\sqrt{z^2 + R^2 - 2zR \cos \vartheta}}$$

$$\begin{aligned}\bar{\phi} &= \frac{-q}{16\pi^2\epsilon_0} \int_0^{2\pi} d\varphi \int_0^\pi d\cos\vartheta \frac{1}{\sqrt{z^2 + R^2 - 2zR\cos\vartheta}} = \\ &= \frac{q}{8\pi\epsilon_0} \frac{1}{zR} \sqrt{z^2 + R^2 - 2zR\cos\vartheta} \Big|_0^\pi = \\ &= \frac{q}{8\pi\epsilon_0} \frac{(z+R) - (z-R)}{zR} = \frac{1}{4\pi\epsilon_0} \frac{q}{z}\end{aligned}$$

Jest to dokładnie wartość potencjału w środku sfery pochodzącego od ładunku q . Pokazaną własność łatwo uogólniamy na bardziej skomplikowane układy ładunków. Potencjał w punkcie p pochodzący od dowolnego rozkładu ładunku jest średnią potencjału liczoną na dowolnej (0 dowolnym promieniu) sferze scentrowanej w punkcie p . Zauważmy, że z własności 2. natychmiast wynika następująca własność:

3. Wewnątrz obszaru pozbawionego ładunku potencjał elektrostatyczny nie może mieć ekstremów. Mogą one występować jedynie na jego brzegu. Wynika to stąd, że jeżeli by istniało minimum (maximum) w pewnym punkcie p wewnątrz obszaru całkowania równania Poissona, wówczas istniałaby sfera otaczająca punkt p na której wszystkie wartości potencjału byłyby większe (mniejsze) - a zatem również ich średnia - od wartości potencjału w punkcie p .

Ta z kolei własność pozwala nam udowodnić twierdzenie 1.

Założmy, że istnieją dwa rozwiązania równania Laplace'a:

$$\nabla^2\phi_1(\vec{r}) = 0 \quad \text{ i } \quad \nabla^2\phi_2(\vec{r}) = 0$$

spełniające te same warunki brzegowe na powierzchni otaczającej rozważany obszar. Ich różnica: $\phi = \phi_1 - \phi_2$ spełnia również równanie Laplace'a, a na brzegu obszaru jest identycznie równa zero. Ponieważ $\phi(\vec{r})$ nie może mieć ekstremów wewnątrz obszaru, musi się w całym obszarze zerować, zatem oba rozwiązania są identyczne.

Podobnie dowodzimy jednoznaczność rozwiązań równania Poissona:

$$\nabla^2\phi_1(\vec{r}) = -\frac{1}{\epsilon_0}\rho \quad \text{ i } \quad \nabla^2\phi_2(\vec{r}) = -\frac{1}{\epsilon_0}\rho$$

ponieważ ich różnica spełnia równanie Laplace'a i zeruje się na brzegu obszaru.

Równanie Poissona (Laplace'a) wykorzystujemy najczęściej w elektrostatyce. W magnetostatyce, w przypadku układów o niskiej symetrii możemy wykorzystać również równania Poissona lub przetworzyć je do postaci dającej prawo Biota- Savarta.

Prawo Biota-Savarta.

Dwa równania zawierające pole magnetyczne

$$\nabla \times \vec{H} = \vec{D} + \vec{j}$$

$$\nabla \cdot \vec{B} = 0$$

przy założeniu niezależności czasowej pól, oraz jednorodności ośrodka przyjmują postać:

$$\nabla \times \vec{H} = \vec{j}$$

$$\nabla \cdot \vec{H} = 0$$

Pole magnetyczne zapisujemy przy pomocy potencjału wektorowego pola

$$\mu_0 \vec{H} = \vec{B} = \text{rot} \vec{A} = \nabla \times \vec{A}$$

Wstawiając ten związek do równania Maxwella otrzymujemy:

$$\nabla \times \nabla \times \vec{A} = \mu_0 \vec{j}$$

Korzystamy ze związku

$$\nabla \times \nabla \times \vec{A} = \nabla(\nabla \cdot \vec{A}) - (\nabla \cdot \nabla)\vec{A} = \text{grad}(\text{div}\vec{A}) - \nabla^2 \vec{A}$$

Uzyskujemy

$$\nabla^2 \vec{A} = -\mu_0 \vec{j} + \text{grad}(\text{div}\vec{A})$$

Niejednoznaczność wyboru czteropotencjału pozwala dodać do niego czterogradient dowolnej funkcji skalarnej:

$$\tilde{\vec{A}}_\mu = A_\mu + \frac{\partial f}{\partial x^\mu}$$

co oznacza dodanie do potencjału wektorowego gradientu funkcji f

$$\tilde{\vec{A}} = \vec{A} + \nabla f$$

Można więc zażądać by

$$\text{div}\tilde{\vec{A}} = 0,$$

ponieważ sprowadzi się to do znalezienia funkcji f spełniającej równanie

$$0 = \nabla \tilde{\vec{A}} = \nabla \vec{A} + \nabla^2 f,$$

czyli równania Poissona

$$\nabla^2 f = -\nabla \cdot \vec{A}$$

Nie musimy go rozwiązywać, wystarczy nam jego rozwiązywalność. Możemy więc narzucić na potencjał wektorowy warunek

$$\text{div}\vec{A} = 0$$

co pozwala odrzucić ostatni wyraz w równaniu określającym źródła pola magnetycznego.

Przyjmuje ono postać:

$$\nabla^2 \vec{A} = -\mu_0 \vec{j}$$

Jest to układ trzech niezależnych równań Poissona na każdą składową potencjału wektorowego i prądu. Analogia z równaniem Poissona dla potencjału skalarnego jest oczywista. Jeżeli założymy podobne warunki brzegowe, czyli znikanie potencjału w nieskończoności, możemy natychmiast napisać jego rozwiązanie:

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \int \frac{\vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}'$$

Dla pola magnetycznego otrzymujemy:

$$\vec{B}(\vec{r}) = \text{rot}\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \nabla_{\vec{r}} \times \int \frac{\vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}'$$

Ponieważ różniczkujemy po innym wektorze niż całkujemy, możemy wejść pod znak całki i otrzymujemy:

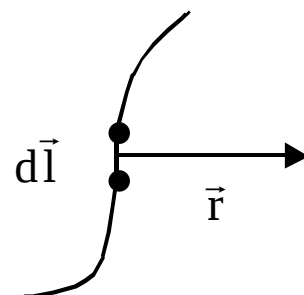
$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi} \int \nabla_{\vec{r}} \times \frac{\vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3\vec{r}' = -\frac{\mu_0}{4\pi} \int \frac{(\vec{r} - \vec{r}') \times \vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|^3} d^3\vec{r}'$$

Jest to prawo Biota-Savarta dla gęstości prądu. Przy założeniu prądu płynącego w cienkim przewodniku można wykonać całkowanie po jego przekroju poprzecznym i wyrażenie na wektor indukcji pola magnetycznego redukuje się do całki jednokrotnej:

$$\vec{B}(\vec{r}) = -\frac{\mu_0}{4\pi} \int \frac{I}{|\vec{r} - \vec{r}'|^3} (\vec{r} - \vec{r}') \times d\vec{r}'$$

Prawo Biota Savarta zazwyczaj zapisuje się w uproszczonej postaci:

$$d\vec{B} = \frac{\mu_0}{4\pi} \frac{I}{r^3} d\vec{l} \times \vec{r}$$



Prawo Biota-Savarta jest prawem całkowym, ale w odróżnieniu od praw Maxwella w postaci całkowitej wygenerowaliśmy je korzystając z postaci różniczkowej równań Maxwella a następnie z zasady superpozycji.

Podsumowanie XI.

- Dla ilustracji związku elektrodynamiki z teorią relatywistyczną, korzystając z transformacji Lorentza wyliczyliśmy rozkład pola magnetycznego wokół przewodnika z prądem korzystając ze znajomości potencjału elektrostatycznego pochodzącego od jednorodnie naładowanego pręta.
- W układach o niskiej symetrii, w których nie jest możliwe wykorzystanie praw Maxwella w postaci całkowitej, stosujemy ich różniczkową postać.
- Rozkładu potencjału elektrostatycznego szukamy rozwiązując równanie Poissona lub Laplace'a.
- Dla uzyskania rozwiązania szczególnego musimy zadać warunki brzegowe (Dirichleta lub von Neumanna) na powierzchni zamkniętej otaczającej interesujący nas obszar.
- Przedyskutowaliśmy jednoznaczność rozwiązań równania Poissona i Laplace'a.
- W przypadku magnetostatyki z równań Maxwella w różniczkowej postaci uzyskujemy równania Poissona dla pola magnetycznego, lub wyrażenia na pole magnetyczne w postaci prawa Biota Savarta.

XII.

Przewodniki i warunki brzegowe na ich powierzchniach.

Powróćmy do dyskusji warunków brzegowych niezbędnych do rozwiązania równania Poissona lub Laplace'a. Z warunkiem brzegowym spotkaliśmy się przy omawianiu potencjału pochodzącego od ładunku punktowego. Wyzaczył on poziom odniesienia względem którego mierzyliśmy energię potencjalną. Założyliśmy, że w nieskończonej odległości od ładunku będącym źródłem pola potencjał jest równy zero. Zerowy potencjał w nieskończoności zawsze możemy przyjąć jako warunek brzegowy w przypadku gdy ładunki będące źródłem pola zawarte są w skończonym obszarze przestrzeni.

Innym obszarem przestrzeni, w którym możemy zadać warunki brzegowe są powierzchnie przewodników. Wynika to z ich podstawowej własności, którą możemy sformułować następująco:

Idealnym przewodnikiem jest substancja zawierająca nieskończony zapas swobodnych elektronów.

Jej konsekwencją są dalsze własności:

1. W całej objętości przewodnika $\vec{E} = 0$. Gdyby tak nie było, swobodne ładunki (a jest ich nieskończony zapas) poruszałyby się w polu elektrycznym aż do jego całkowitej kompensacji.
2. Na zewnętrznej powierzchni przewodnika pole elektryczne jest do niej prostopadłe. Składowa pola elektrycznego równoległa do powierzchni przewodnika wywołuje ruch elektronów i jej kompensację.
3. $\rho(\vec{r}) = 0$ w całej objętości przewodnika. Jest to konsekwencja zerowania się pola elektrycznego i (trzeciego) prawa Maxwella $\rho = \epsilon_0 \nabla \cdot \vec{E}$.
4. Nie skompensowany ładunek może wystąpić jedynie na powierzchni przewodnika. Tam występuje skok (nieciągłość) składowej pola elektrycznego prostopadłej do powierzchni przewodnika.
5. Potencjał wewnątrz jednolitego obszaru przewodnika i na jego powierzchni jest stały, ponieważ dla dwóch punktów $\vec{r}_1$ i $\vec{r}_2$ dowolnie wybranych w takim obszarze:

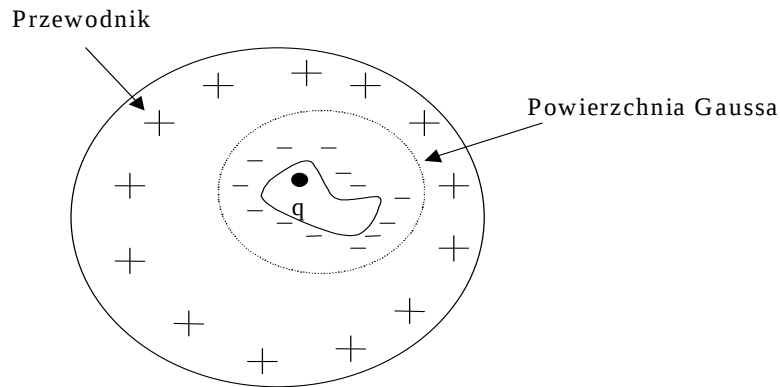
$$V(\vec{r}_2) = V(\vec{r}_1) + \int_{\vec{r}_1}^{\vec{r}_2} \vec{E} \cdot d\vec{r} = V(\vec{r}_1)$$

Wymienione własności przewodników pozwalają na ustalenie warunków brzegowych na ich powierzchniach.

Ładunki indukowane.

Jeżeli kawałek przewodnika umieścimy w zewnętrznym polu elektrycznym, np. wytworzonym przez ładunek znajdujący się w pewnej od niego odległości, wówczas wewnątrz przewodnika następuje taka redystrybucja ładunku, aby wyzerowała się równoległa do powierzchni przewodnika składowa pola elektrycznego. Taka kompensacja jest wynikiem pojawienia się dodatkowego pola elektrycznego, którego źródłem jest wytworzony na skutek tejże redystrybucji ładunek powierzchniowy nazywany ładunkiem indukowanym. Zjawisko to ma czasem interesujące konsekwencje. Rozważmy własność stanowiącą ciekawostkę, która po przeanalizowaniu jest dość oczywista, ale na pierwszy rzut oka nieoczekiwana.

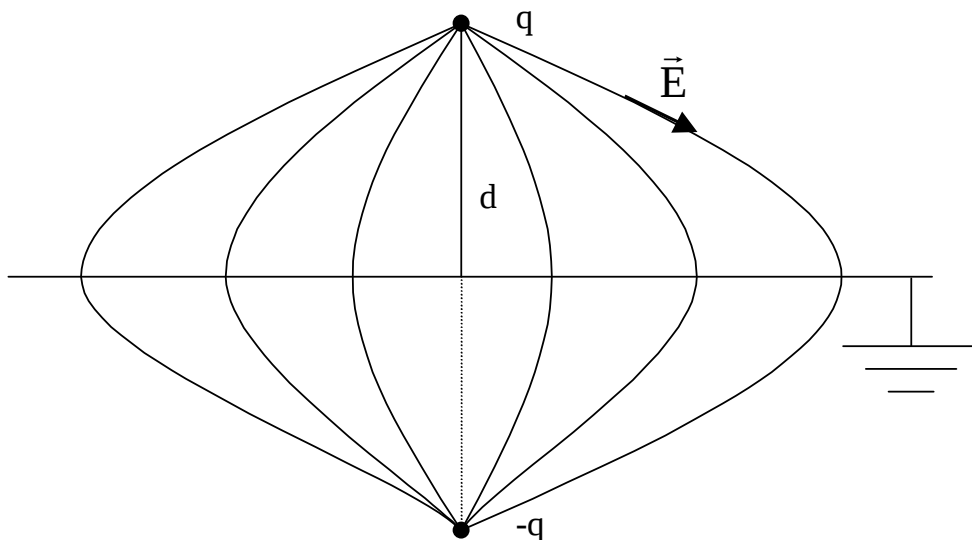
We wnętrzu kuli zbudowanej z przewodnika znajduje się nie przewodząca wnęka o nieregularnym kształcie, w której znajduje się dodatni ładunek q a sam przewodnik nie jest naładowany. Rysunek poniżej ilustruje przekrój przez kulę i wnękę.



Wewnątrz wnęki pojawia się niezerowe pole elektryczne. Wywołuje ono redystrybucję ładunku wewnątrz przewodnika. Na wewnętrznej powierzchni (granica przewodnik - wnęka) indukuje się ładunek przeciwnego znaku do q , ekranując go całkowicie, to znaczy tak, by wewnątrz przewodnika pole elektryczne było równe zero. Jeżeli teraz otoczmy wnękę powierzchnią zamkniętą, zaznaczoną na rysunku linią przerywaną, to zgodnie z prawem Gaussa całkowity ładunek objęty tą powierzchnią musi się zerować. Wielkość ładunku indukowanego musi być więc równa dokładnie $-q$. Ponieważ założyliśmy neutralność przewodnika, obsadzeniu powierzchni wewnętrznej ładunkiem $-q$ musi towarzyszyć obsadzenie powierzchni zewnętrznej ładunkiem q . Tym razem rozkład ładunku na powierzchni przewodnika jest równomierny a informacja o kształcie jamy zostaje całkowicie ukryta przez przewodnik.

Pole elektrostatyczne w obecności przewodników. Metoda obrazów.

Jednoznaczność rozwiązań równania Poissona dla zadanych warunków brzegowych pozwala w ułatwiony sposób liczyć pole elektrostatyczne w obecności przewodników. Metoda rachunkowa nosi nazwę metody obrazów.



Założmy, że pojedynczy punktowy ładunek q umieszczony jest w odległości d od uziemionej płaszczyzny przewodzącej. Znajdźmy rozkład potencjału elektrostatycznego. Jeżeli wybierzemy układ współrzędnych z osiami x, y położonymi na przewodzącej płaszczyźnie, ładunek położony jest w punkcie $(0, 0, d)$. Musimy rozwiązać równanie Poissona w górnej

półprzestrzeni z warunkiem brzegowym: $\phi(\vec{r}) = 0$ we wszystkich punktach przewodzącej płaszczyzny, oraz w nieskończonej odległości od ładunku q . Zauważmy, że taki sam warunek będzie spełniony w problemie, w którym nie mamy przewodzącej półpłaszczyzny, ale za to mamy drugi ładunek $-q$ umieszczony w punkcie $(0,0,-d)$, a więc jak gdyby zwierciadlane odbicie w płaszczyźnie przewodnika.

W rezultacie w punkcie (x,y,z) potencjał dany jest wyrażeniem:

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{\sqrt{x^2 + y^2 + (z-d)^2}} - \frac{q}{\sqrt{x^2 + y^2 + (z+d)^2}} \right)$$

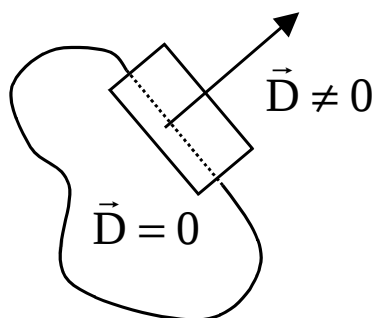
Oczywiście potencjał ten spełnia w górnej półpłaszczyźnie właściwe równanie Poissona, a jego wartość na płaszczyźnie „xy” (podstawiamy $z=0$) jest stała i równa 0. Dzięki twierdzeniu o jednoznaczności sposób dojścia do rozwiązania równania Poissona nie jest istotny. Trick pozwalający na uwzględnienie obecności przewodników w przestrzeni przez dodanie do problemu dodatkowych ładunków pilnujących właściwych warunków brzegowych nosi nazwę metody obrazów. Przedstawiona metoda po prostym uogólnieniu pozwala na wyliczenie rozkładu pola pochodzącego od dowolnego układu ładunków w obecności przewodzących płaszczyzn. Metodę obrazów możemy również zastosować przy liczeniu rozkładu potencjału w obecności przewodników w kształcie kuli lub walca.

Obraz a ładunek indukowany.

Obraz ładunku pierwotnego nie jest oczywiście żadnym realnym ładunkiem, jego obecność jest jedynie konstrukcją matematyczną pozwalającą na uwzględnienie warunków brzegowych. Jednakże na powierzchni przewodnika pojawia się realny ładunek indukowany, stanowiący wraz z ładunkiem pierwotnym źródło pola. Znając (np. uzyskany metodą obrazów) rozkład potencjału w przestrzeni możemy korzystając z równania Poissona wyliczyć faktyczny rozkład ładunków:

$$\rho(\vec{r}) = -\epsilon_0 \nabla^2 \phi(\vec{r}).$$

Należy jednak pamiętać, że równanie Poissona nie odróżnia ładunków indukowanych od pierwotnych. Sam indukowany ładunek pojawia się na powierzchni przewodnika w postaci ładunku powierzchniowego, a więc posiada nieskończenie dużą gęstość w kierunku prostopadłym do powierzchni. Nie możemy go wyliczyć różniczkując potencjał lub pole elektryczne. Możemy go jednak wyliczyć korzystając z uzyskanych poprzednio własności



przewodnika.

Wewnątrz przewodnika pole elektryczne się zeruje. Jeżeli na zewnątrz przewodnika znajduje się niezerowe pole elektryczne, jest ono prostopadłe do jego powierzchni. Wybierzmy niewielki obszar powierzchni przewodnika i otoczmy go prostopadłościenną zamkniętą powierzchnią. Zgodnie z prawem Maxwella (3):

$$\oint_{\Sigma} \vec{D} d\vec{s} = q.$$

Ponieważ całka po powierzchni dolnej i bocznych się zeruje dla elementarnie małej powierzchni, możemy napisać:

$$\vec{D} d\vec{s} = dq = \sigma ds,$$

gdzie przez σ oznaczyliśmy gęstość powierzchniową ładunku. Oznaczając przez $D_{\perp}$ prostopadłą do powierzchni przewodnika składową pola elektrycznego otrzymujemy:

$$D_{\perp} ds = \sigma ds$$

skąd

$$D_{\perp} = \sigma.$$

Ponieważ wyliczyliśmy rozkład potencjału elektrostatycznego wywołany przez ładunek umieszczony ponad przewodzącą płaszczyznę, sprawdźmy jaki ładunek zostanie wyindukowany na tej powierzchni. Skorzystamy ze związku:

$$\vec{D} = -\epsilon_0 \nabla \phi,$$

do którego wstawiamy wyrażenie na potencjał:

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{\sqrt{x^2 + y^2 + (z-d)^2}} - \frac{q}{\sqrt{x^2 + y^2 + (z+d)^2}} \right)$$

$$\vec{D} = \frac{q}{4\pi} \left(\frac{(x, y, z-d)}{\left(x^2 + y^2 + (z-d)^2\right)^{\frac{3}{2}}} - \frac{(x, y, z+d)}{\left(x^2 + y^2 + (z+d)^2\right)^{\frac{3}{2}}} \right)$$

Na powierzchni przewodnika ($z=0$):

$$\begin{aligned} \vec{D} &= \frac{q}{4\pi} \left(\frac{(x, y, -d)}{\left(x^2 + y^2 + d^2\right)^{\frac{3}{2}}} - \frac{(x, y, d)}{\left(x^2 + y^2 + d^2\right)^{\frac{3}{2}}} \right) = \\ &= -\frac{qd}{2\pi} \frac{(0,0,1)}{\left(x^2 + y^2 + d^2\right)^{\frac{3}{2}}}. \end{aligned}$$

Ponieważ

$$D_{\perp} = D_z,$$

na przewodniku indukuje się ładunek o gęstości powierzchniowej:

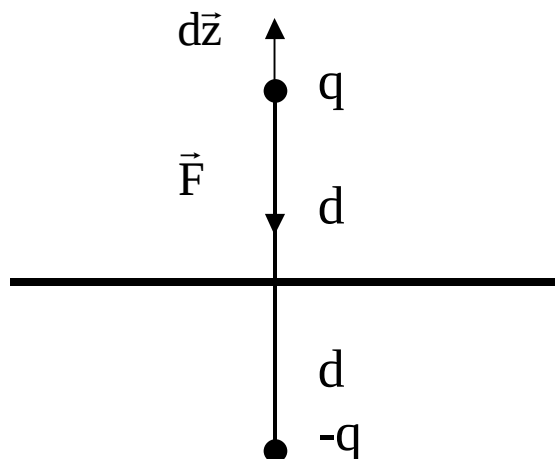
$$\sigma = D_{\perp} = -\frac{qd}{2\pi} \frac{1}{\left(x^2 + y^2 + d^2\right)^{\frac{3}{2}}}$$

Wyliczyliśmy rozkład gęstości ładunku indukowanego. Możemy sprawdzić jak duży jest całkowity ładunek indukowany:

$$\begin{aligned} q^{\text{ind}} &= \int \sigma(x, y) dx dy = \int_0^{2\pi} \int_0^{\infty} dr r \sigma(r) = -\frac{qd}{2\pi} \int_0^{2\pi} d\phi \int_0^{\infty} dr r \frac{1}{\left(r^2 + d^2\right)^{\frac{3}{2}}} \\ &= -qd \frac{1}{2} \int_0^{\infty} d(r^2) \frac{1}{\left(r^2 + d^2\right)^{\frac{3}{2}}} = -qd \frac{1}{2} \int_0^{\infty} dx \frac{1}{\left(x + d^2\right)^{\frac{3}{2}}} = qd \frac{1}{\left(x + d^2\right)^{\frac{1}{2}}} \Big|_0^{\infty} = -q \end{aligned}$$

Okazuje się, że całkowity indukowany ładunek jest równy ładunkowi odpowiedniego obrazu. Można się tego było spodziewać mając przed oczami trzecie prawo Maxwella w postaci całkowej.

Jest oczywiste, że ładunek indukowany wytwarza pole elektrostatyczne, które działa na ładunek wywołujący zjawisko. Zastanówmy się teraz jaka siła będzie działać na ładunek ze strony obrazu i jaka będzie energia potencjalna ładunku w zadanej odległości od przewodnika. Wynik okaże się małą niespodzianką.



Potencjał indukowany w punkcie, w którym znajduje się ładunek q liczymy korzystając z metody obrazów:

$$V^{\text{ind}}(d) = \frac{-q}{4\pi\epsilon_0 z} \Big|_{z=2d} = \frac{-q}{8\pi\epsilon_0 d}$$

Energię potencjalną ładunku w polu liczymy zazwyczaj mnożąc ładunek przez wartość potencjału. Jest to słuszne zawsze z wyjątkiem przypadku potencjału indukowanego, w którym postępujemy inaczej.

Siła działająca na ładunek jest równoległa do kierunku „ z ”, a jej wartość:

$$\begin{aligned} F = -qE = -q \frac{\partial}{\partial z} V^{\text{ind}}(z) \Big|_{z=2d} &= -q \frac{\partial}{\partial z} \frac{-q}{4\pi\epsilon_0 z} \Big|_{z=2d} = \frac{-q^2}{4\pi\epsilon_0 z^2} \Big|_{z=2d} \\ &= -\frac{q^2}{4\pi\epsilon_0} \frac{1}{4d^2} = -\frac{q^2}{16\pi\epsilon_0} \frac{1}{d^2} \end{aligned}$$

Praca potrzebna do przeniesienia ładunku z punktu d do nieskończoności stanowi jego energię potencjalną:

$$E_p = \int_d^\infty \frac{-q^2}{16\pi\epsilon_0} \frac{1}{z^2} dz = \frac{q^2}{16\pi\epsilon_0} \frac{1}{z} \Big|_d^\infty = \frac{-q^2}{16\pi\epsilon_0} \frac{1}{d}$$

Energia potencjalna ładunku jest dwa razy mniejsza niż wynikałoby to z wielkości potencjału indukowanego. Oddziaływanie z ładunkiem obrazu ma charakter samooddziaływania.

Podczas przesuwania ładunku odsuwa się równocześnie obraz.

Wiemy zatem jak radzić sobie w przypadku obecności przewodników.

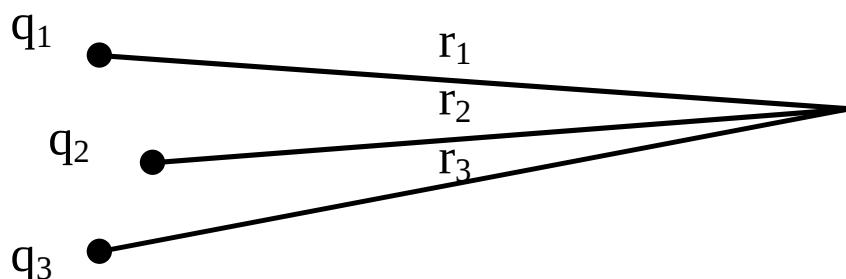
Przejdziemy teraz do omówienia zjawisk pojawiających się w ośrodkach i na ich granicach.

Zanim jednak przejdziemy do tego tematu, przedstawię pewną metodę rachunkową, wykorzystywaną w elektrostatyce, do której elementów odwołam się dyskutując pola elektryczne w ośrodkach.

Rozwinięcie multipolowe

Czasem się zdarza, że źródłem pola jest układ ładunków znajdujących się daleko od obszaru, w którym liczymy potencjał pola.

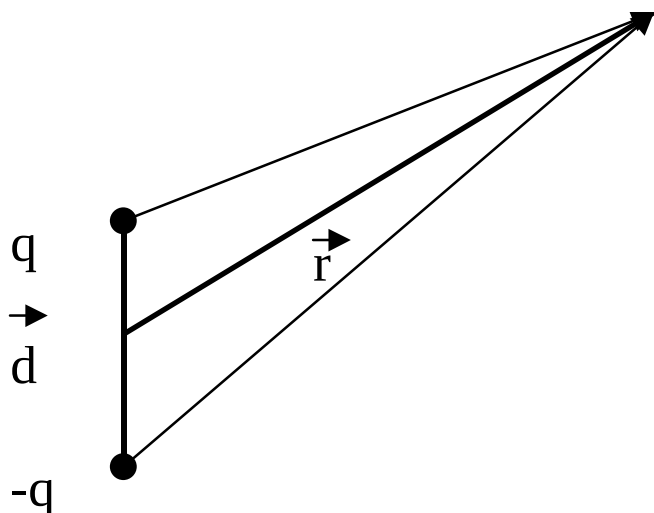
Wówczas w pierwszym przybliżeniu można zsumować ładunki, uśrednić odległości i uzyskać



potencjał wypadkowy:

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \frac{\sum_i q_i}{\bar{r}}$$

Przybliżenie takie jest sensowne w zasadzie tylko wtedy gdy ładunek sumaryczny jest różny od zera. Jednakże potencjał układu ładunków nie musi się zerować również w przypadku gdy ładunek całkowity jest zerowy. Rozważmy dwa ładunki jednakowej wielkości ale przeciwnego znaku q i $-q$, umieszczone w odległości d od siebie. Układ taki nosi nazwę dipola. Policzmy potencjał wytworzony w odległości r od środka odcinka d .



$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{\left| \vec{r} - \frac{\vec{d}}{2} \right|} + \frac{-q}{\left| \vec{r} + \frac{\vec{d}}{2} \right|} \right) =$$

$$= \frac{1}{4\pi\epsilon_0} \left(\frac{q}{\sqrt{\vec{r}^2 - \vec{d} \cdot \vec{r} + \frac{\vec{d}^2}{4}}} - \frac{q}{\sqrt{\vec{r}^2 + \vec{d} \cdot \vec{r} + \frac{\vec{d}^2}{4}}} \right)$$

Przy założeniu $r \gg d$, możemy dokonać rozwinięcia w szereg względem d/r i ograniczyć się do wyrazów pierwszego rzędu. Otrzymujemy:

$$\begin{aligned} V(\vec{r}) &= \frac{1}{4\pi\epsilon_0} \left(\frac{q}{r\sqrt{1 - \frac{\vec{r}\vec{d}}{r^2}}} - \frac{q}{r\sqrt{1 + \frac{\vec{r}\vec{d}}{r^2}}} \right) = \\ &= \frac{1}{4\pi\epsilon_0} \frac{q}{r} \left(1 + \frac{\vec{r}\vec{d}}{2r^2} - 1 + \frac{\vec{r}\vec{d}}{2r^2} \right) = \frac{1}{4\pi\epsilon_0} \frac{q\vec{r}\vec{d}}{r^3} = \frac{1}{4\pi\epsilon_0} \frac{qd \cos \theta}{r^2} \end{aligned}$$

Potencjał nie jest zerowy, ale spada z odległością szybciej niż potencjał kulombowski, jest proporcjonalny do odwrotności kwadratu odległości.

Układ złożony z czterech ładunków ułożonych tak, by zniknął również potencjał dipolowy, kwadrupol daje potencjał spadający z 3 potęgą promienia. Oczywiście można skonstruować multipole wyższych rzędów, a potencjały pochodzące od nich spadają z kolejnymi wyższymi potęgami odwrotności promienia.

Wykorzystajmy te zależności w trochę ogólniejszej dyskusji. Rozważmy potencjał pochodzący od ciągłego rozkładu ładunków zadany gęstością $\rho(\vec{r})$. Zgodnie z poprzednio uzyskanymi z zasady superpozycji wzorami potencjał od niego pochodzący

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r'$$

Jeżeli badamy potencjał w dużej odległości od źródeł, możemy oczekiwać, że potencjał ten uda nam się rozwinąć w szereg multipolowy, w którym kolejne wyrazy będą zanikać zgodnie z kolejnymi potęgami odwrotności odległości. W tym celu posłużymy się rozwinięciem na wielomiany Legendre'a:

$$\frac{1}{|\vec{r} - \vec{r}'|} = \frac{1}{r} \sum_{n=0}^{\infty} \left(\frac{r'}{r} \right)^n P_n \left(\frac{\vec{r}\vec{r}'}{rr'} \right) = \frac{1}{r} \sum_{n=0}^{\infty} \left(\frac{r'}{r} \right)^n P_n(\cos \vartheta)$$

gdzie ϑ jest kątem pomiędzy oboma wektorami. Rozwinięcie to wstawiamy pod całkę i otrzymujemy:

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \int r'^n \rho(\vec{r}') P_n \left(\frac{\vec{r}\vec{r}'}{rr'} \right) d^3r',$$

Pierwszy wyraz, po podstawieniu $P_0(x) = 1$

daje potencjał monopolowy:

$$V_{\text{mon}}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \frac{Q}{r}$$

gdzie $Q = \int \rho(\vec{r}') d^3r'$ jest ładunkiem całkowitym.

Drugi wyraz rozwinięcia jest potencjałem dipolowym, uzyskujemy go podstawiając $P_1(x) = x$

$$\begin{aligned} V_{\text{dip}}(\vec{r}) &= \frac{1}{4\pi\epsilon_0} \frac{1}{r^2} \int r' \rho(\vec{r}') \frac{\vec{r}\vec{r}'}{rr'} d^3r' = \frac{1}{4\pi\epsilon_0} \frac{1}{r^2} \int \rho(\vec{r}') \frac{\vec{r}\vec{r}'}{r} d^3r' \\ &= \frac{1}{4\pi\epsilon_0} \frac{1}{r^2} \int \rho(\vec{r}') r' \cos \vartheta d^3r' \end{aligned}$$

Można go zapisać nieco inaczej wyciągając $\vec{r}$ przed całkę:

$$V_{\text{dip}}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \frac{\vec{r}}{r^3} \int \rho(\vec{r}') \vec{r}' d^3r' = \frac{1}{4\pi\epsilon_0} \frac{\vec{p}}{r^2} \frac{\vec{r}}{r},$$

gdzie

$$\vec{p} = \int \rho(\vec{r}') \vec{r}' d^3r'$$

nazywamy momentem dipolowym rozkładu ładunku. W przypadku ładunków punktowych całka przechodzi w sumę:

$$\vec{p} = \sum_i q_i \vec{r}_i,$$

co dla dwóch przeciwnych ładunków ułożonych w odległości $\vec{d}$ daje $\vec{p} = q\vec{d}$.

Podobnie możemy postępować dalej i uzyskać kolejno potencjał kwadrupolowy i potencjały wyższych rzędów.

Podsumowanie XII.

- Naturalne warunki brzegowe możemy zadać w dwóch przypadkach.
- ♦ W nieskończoności - jeżeli ładunek będący źródłem pola rozłożony jest w ograniczonym obszarze przestrzennym (żądamy zerowania się potencjału w nieskończoności).
- ♦ Na powierzchni przewodnika gdzie potencjał jest stały.
- ♦ Na powierzchni przewodnika pojawiają się ładunki indukowane, które korygują rozkłady potencjału.
- Spełnienie warunku brzegowego można zasymulować stosując tzw. metodę obrazów.
- Obrazy ładunku symulują działanie ładunku indukowanego, wielkości ładunku obrazu i ładunku indukowanego są jednakowe, różne jest ich położenie. Obraz umieszczamy wewnątrz przewodnika, a ładunek indukowany pojawia się na jego powierzchni.
- Przedyskutowaliśmy rozwinięcie multipolowe potencjału pochodzącego od ciągłego rozkładu ładunku.

XIII.

Pola elektryczne i magnetyczne w ośrodkach, zależności pomiędzy $\vec{E}$ i $\vec{D}$ oraz $\vec{H}$ i $\vec{B}$.

W prowadzonych rozważaniach pojawiały się w różnych wyrażeniach zamiennie wektory: natężenia pola elektrycznego $\vec{E}$ i indukcji elektrycznej $\vec{D}$ oraz natężenia pola magnetycznego $\vec{H}$ i indukcji magnetycznej $\vec{B}$. Nie wprowadzałem rozróżnienia pomiędzy nimi, zakładając, że jest ono znane z wykładu fizyki ogólnej. Jednakże w celu uporządkowania pojęć, chciałbym zwrócić uwagę na związki i różnice pomiędzy odpowiednimi parami wektorów. W ośrodkach izotropowych (próżni) łączą je proste związki, które wykorzystywałem:

$$\vec{D} = \epsilon \epsilon_0 \vec{E}$$

$$\vec{B} = \mu \mu_0 \vec{H}$$

$\epsilon(\mu)$ są bezwymiarowymi stałymi nazywanymi przenikalnością dielektryczną (magnetyczną) ośrodka. W próżni i ośrodkach izotropowych wektory $\vec{E}$ i $\vec{D}$ ($\vec{H}$ i $\vec{B}$) są odpowiednio równoległe. W ośrodkach anizotropowych związki te pozostają liniowe, ale bardziej złożone:

$$D^i = \epsilon_0 \sum_j \epsilon_{ij} E^j$$

$$B^i = \mu_0 \sum_j \mu_{ij} H^j$$

Przenikalności dielektryczna i magnetyczna stają się wielkościami tensorowymi. Wektory $\vec{E}$ i $\vec{D}$ ($\vec{H}$ i $\vec{B}$) przestają być równoległe.

Istota wektorów pól.

Pozostańmy jednak przy ośrodkach izotropowych (próżni). Wektory $\vec{E}$ i $\vec{D}$ ($\vec{H}$ i $\vec{B}$) wyrażane są w innych jednostkach. Wielkości $\epsilon_0(\mu_0)$ noszące nazwy przenikalności elektrycznej (magnetycznej) próżni są w układzie SI mianowane i odpowiedzialne za zmianę jednostek.

By zauważyć różnice pomiędzy parami wektorów musimy przyjrzeć się wyrażeniom w których występują. Pierwszym z nich jest wzór na siłę Lorentza:

$$\vec{F} = e\vec{E} + e\vec{V} \times \vec{B}$$

Pola $\vec{E}$ i $\vec{B}$ związane są z siłami działającymi na ładunek w polu elektrycznym i magnetycznym. Pozostałe dwa wektory pojawiają się w równaniach:

$$\nabla \cdot \vec{D} = \rho \quad \text{ i } \quad \nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}$$

w których występują źródła pól gęstość ładunku lub prąd. Równania te decydują o jednostkach, w których wyrażane są odpowiednie wektory i ostatecznie determinują przenikalności dielektryczne (magnetyczne) próżni.

Polaryzacja dielektryka.

Zastanówmy się teraz co się dzieje w dielektrykach umieszczonych w polu elektrycznym. Naturalnie ponieważ dielektryki nie zawierają ładunków swobodnych, nie nastąpi w nich przepływ ładunku i całkowita kompensacja pola jak miało miejsce w przewodnikach. Pole będzie wnikać do wnętrza dielektryka i oddziaływać z atomami czy cząsteczkami z których zbudowany jest materiał. Cząsteczki mogą wykazywać spontaniczną polaryzację, posiadać moment dipolowy bez pola elektrycznego, przykładem takiej cząsteczki jest H_2O , posiadająca asymetryczną budowę. Atom tlenu i dwa atomy wodoru tworzą trójkąt a chmura elektronowa jest przesunięta bardziej w kierunku atomu tlenu.

$$\vec{p} = q\vec{d}$$

Inne cząsteczki albo atomy mogą nie wykazywać momentu dipolowego bez pola elektrycznego ale pod jego wpływem chmury elektronowe przesuwają się względem jąder, atomy ulegają polaryzacji. Indukowany moment dipolowy jest proporcjonalny do natężenia pola elektrycznego $\vec{p} = \alpha \vec{E}$.

Wypadkowa siła działająca na dipol w jednorodnym polu elektrycznym zeruje się.

$$\vec{F} = \pm q\vec{E}$$

Nie zeruje się natomiast moment siły

$$\vec{N} = \vec{d} \times \vec{F} = \vec{p} \times \vec{E}.$$

Dipole będą ustawiać się zgodnie z polem elektrycznym. Miara stopnia uporządkowania dipoli jest tzw. polaryzacja elektryczna P definiowana jako moment dipolowy jednostki objętości dielektryka.

Polaryzacja dielektryka jest źródłem dodatkowego pola elektrycznego. Spróbujmy policzyć jego wielkość. Potencjał pochodzący od pojedynczego dipola umieszczonego w początku układu dany jest wyrażeniem:

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \frac{\vec{r} \cdot \vec{p}}{r^3}.$$

Potencjał pochodzący od ciągłego rozkładu gęstości dipoli zebranych w objętości V uzyskamy zastępując pojedynczy moment dipolowy rozkładem ciągłym:

$$\vec{p} \rightarrow \vec{P}(\vec{r})dV$$

i wykonując całkowanie po całej objętości:

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_V \frac{(\vec{r} - \vec{r}') \cdot \vec{P}(\vec{r}')}{|\vec{r} - \vec{r}'|^3} d^3r'.$$

Przekształcamy ten wzór korzystając ze związku:

$$\frac{\vec{r}}{r^3} = -\nabla \frac{1}{r}.$$

Otrzymujemy:

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int_V \vec{P}(\vec{r}') \nabla_{r'} \frac{1}{|\vec{r} - \vec{r}'|} d^3r'$$

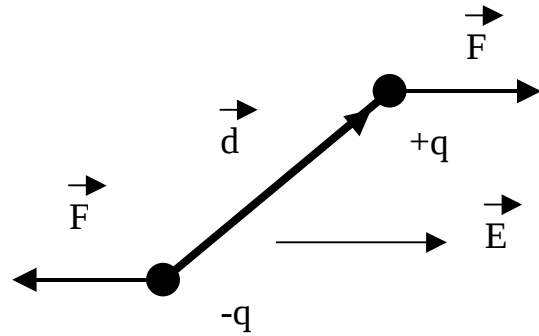
Ten wzór przekształcamy korzystając ze wzoru na pochodną iloczynu

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \left(\int_V \nabla_{r'} \frac{\vec{P}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' - \int_V \frac{1}{|\vec{r} - \vec{r}'|} \nabla_{r'} \vec{P}(\vec{r}') d^3r' \right).$$

Tutaj przekształcamy pierwszy składnik korzystając z prawa Gaussa:

$$\begin{aligned} V(\vec{r}) &= \frac{1}{4\pi\epsilon_0} \left(\oint_{\Sigma} \frac{\vec{P}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{\sigma} - \int_V \frac{1}{|\vec{r} - \vec{r}'|} \nabla_{r'} \vec{P}(\vec{r}') d^3r' \right) = \\ &= \frac{1}{4\pi\epsilon_0} \left(\oint_{\Sigma} \frac{\sigma_{zw}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\sigma - \int_V \frac{\rho_{zw}(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r' \right). \end{aligned}$$

Pierwszy wyraz ma sens rozkładu gęstości ładunku powierzchniowego



$$\sigma_{zw} = \vec{P} \cdot \vec{n}$$

a drugi rozkładu gęstości ładunku przestrzennego

$$\rho_{zw}(\vec{r}) = -\nabla \vec{P}(\vec{r})$$

Oba typy gęstości ładunku będziemy nazywać ładunkiem związanym, żeby podkreślić, że nie nastąpił tutaj przepływ ładunku lecz jedynie reorientacja dipoli.

Wypadkowe pole elektryczne.

Założmy, że w dielektryku mamy zadany rozkład ładunku $\rho_{sw}(\vec{r})$. Dla odróżnienia od ładunku związanego nazwiemy go ładunkiem swobodnym. Jest on źródłem pola elektrycznego, ono z kolei powoduje polaryzację dielektryka i pojawia się dodatkowy ładunek związany. W rezultacie całkowity rozkład ładunku w próżni jest:

$$\rho(\vec{r}) = \rho_{sw}(\vec{r}) + \rho_{zw}(\vec{r})$$

Korzystając z prawa Maxwella dla próżni mamy:

$$\epsilon_0 \nabla \vec{E} = \rho = \rho_{sw} + \rho_{zw} = \rho_{sw} - \nabla \vec{P}$$

zatem

$$\nabla(\epsilon_0 \vec{E} + \vec{P}) = \rho_{sw}$$

a ponieważ

$$\nabla \vec{D} = \rho_{sw}$$

otrzymujemy:

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P}$$

Wektor indukcji elektrycznej D jest powiększany w stosunku do odpowiedniej wartości w próżni o polaryzację ośrodka.

Dielektryki liniowe.

Dla wielu materiałów polaryzacja dielektryka jest proporcjonalna do natężenia pola elektrycznego (w rezultacie momentów sił działających na dipole). Takie materiały nazywamy dielektrykami liniowymi.

$$\vec{P} = \chi_e \epsilon_0 \vec{E}$$

gdzie bezwymiarowy współczynnik proporcjonalności χ_e nazywamy podatnością elektryczną ośrodka. Wtedy

$$\vec{D} = \epsilon_0 \vec{E} + \vec{P} = \epsilon_0 \vec{E} + \chi_e \epsilon_0 \vec{E} = (1 + \chi_e) \epsilon_0 \vec{E} = \epsilon \epsilon_0 \vec{E}$$

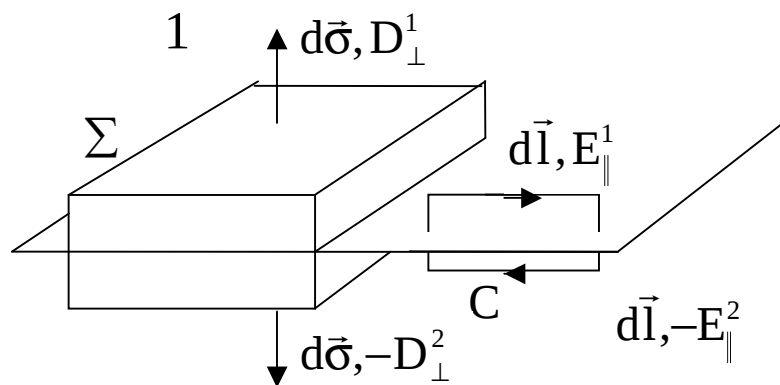
Współczynnik proporcjonalności nazywamy przenikalnością dielektryczną ośrodka, a

bezwymiarowe $\epsilon = (1 + \chi_e)$

przenikalnością względną.

Granice ośrodków.

Zastanówmy się teraz jak zachowują się wektory pól elektrycznego i magnetycznego na granicy dwóch ośrodków jednorodnych. W tym celu skorzystamy z odpowiednich równań Maxwella w postaci całkowej:



$$\oiint_{\Sigma} \vec{D} \cdot d\vec{\sigma} = q$$

$$\oint_C \vec{E} \cdot d\vec{l} = \frac{d}{dt} \iint_S \vec{B} d\vec{\sigma}$$

Pierwsze z nich przy założeniu braku swobodnego ładunku powierzchniowego na granicy ośrodka daje ciągłość prostopadłej do powierzchni składowej pola indukcji elektrycznej $D_{\perp}$:

$$D_{\perp}^1 d\sigma - D_{\perp}^2 d\sigma = 0.$$

Drugie w przypadku pól statycznych zapewnia ciągłość równoległej do granicy ośrodków składowej natężenia pola elektrycznego $E_{\parallel}$:

$$E_{\parallel}^1 dl - E_{\parallel}^2 dl = 0.$$

Zatem:

$$E_{\parallel}^1 = E_{\parallel}^2 \quad \text{i} \quad D_{\perp}^1 = D_{\perp}^2$$

drugi związek można przepisać w postaci

$$\epsilon_1 E_{\perp}^1 = \epsilon_2 E_{\perp}^2, \text{ skąd}$$

$$E_{\perp}^1 = \frac{\epsilon_2}{\epsilon_1} E_{\perp}^2.$$

Dzięki temu:

$$\operatorname{tg} \alpha_1 = \frac{E_{\perp}^1}{E_{\parallel}^1} = \frac{\epsilon_2}{\epsilon_1} \frac{E_{\perp}^2}{E_{\parallel}^2} = \frac{\epsilon_2}{\epsilon_1} \operatorname{tg} \alpha_2$$

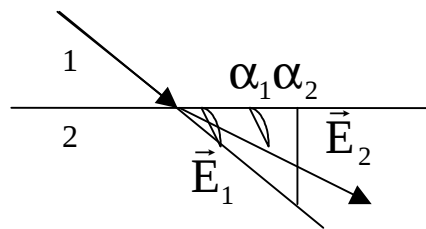
$$\text{lub } \epsilon_1 \operatorname{tg} \alpha_1 = \epsilon_2 \operatorname{tg} \alpha_2$$

Podobnie dla pola magnetycznego z równań:

$$\oiint_{\Sigma} \vec{B} \cdot d\vec{\sigma} = 0 \quad \text{i} \quad \oint_C \vec{H} \cdot d\vec{l} = \frac{d}{dt} \iint_S \vec{D} d\vec{\sigma} + I$$

uzyskujemy ciągłość składowych $B_{\perp}$ i $H_{\parallel}$ na granicy ośrodków i podobną zależność:

$$\mu_1 \operatorname{tg} \alpha_1 = \mu_2 \operatorname{tg} \alpha_2$$



Zmienne pole elektromagnetyczne.

W dotychczasowych rozważaniach dyskutowaliśmy pola elektryczne i magnetyczne, których źródłami były prądy i rozkłady ładunków. Wystarczy jednakże spojrzeć na równanie Maxwella, aby przekonać się, że źródłami pól mogą być one same.

$$1. \nabla \times \vec{E} = -\dot{\vec{B}}$$

$$2. \nabla \cdot \vec{B} = 0$$

$$3. \nabla \cdot \vec{D} = \rho$$

$$4. \nabla \times \vec{H} = \dot{\vec{D}} + \vec{j}$$

Przypomnijmy ponadto związek pól elektrycznego i magnetycznego z potencjałem skalarnym i wektorowym:

$$\vec{E} = -\nabla\phi - \frac{\partial}{\partial t} \vec{A}$$

$$\vec{B} = \nabla \times \vec{A}$$

Liczymy rotacje obu stron ostatniego równania

$$\nabla \times (\nabla \times \vec{A}) = \operatorname{rot} \vec{B} = \mu_0 \nabla \times \vec{H} = \mu_0 (\dot{\vec{D}} + \vec{j}) = \mu_0 \epsilon_0 \dot{\vec{E}} + \mu_0 \vec{j}$$

prawą stronę równania przekształcamy dalej zastępując pole E potencjałami:

$$= -\frac{1}{c^2} \frac{\partial}{\partial t} \left(\nabla \varphi + \frac{\partial}{\partial t} \vec{A} \right) + \mu_0 \vec{j} = -\frac{1}{c^2} \frac{\partial}{\partial t} \nabla \varphi - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} + \mu_0 \vec{j}$$

Natomiast do lewej strony stosujemy wzór znany z analizy wektorowej :

$$\nabla \times (\nabla \times \vec{A}) = \text{grad div} \vec{A} - \nabla^2 \vec{A}$$

Po połączeniu uzyskujemy

$$\text{grad div} \vec{A} - \nabla^2 \vec{A} = -\frac{1}{c^2} \nabla \frac{\partial}{\partial t} \varphi - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} + \mu_0 \vec{j}$$

co po przegrupowaniu daje:

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} = \text{grad div} \vec{A} + \frac{1}{c^2} \nabla \frac{\partial}{\partial t} \varphi - \mu_0 \vec{j} = \nabla \left(\text{div} \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \varphi \right) - \mu_0 \vec{j}$$

Wyraz po prawej stronie równania stanowi czterodivergencję czteropotencjału. Korzystając z jego niejednoznaczności można założyć

$$\text{div} \vec{A} + \frac{1}{c^2} \frac{\partial}{\partial t} \varphi = 0$$

Jest to tzw. cechowanie Lorentza. Dostajemy na potencjał wektorowy równanie:

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} = -\mu_0 \vec{j}$$

Podobnie uzyskujemy równanie dla potencjału skalarnego:

$$\frac{\rho}{\epsilon_0} = \nabla \cdot \vec{E} = -\nabla^2 \varphi - \frac{\partial}{\partial t} \nabla \cdot \vec{A}$$

które korzystając z cechowania Lorentza przekształcamy w:

$$\frac{\rho}{\epsilon_0} = -\nabla^2 \varphi - \frac{\partial}{\partial t} \text{div} \vec{A} = -\nabla^2 \varphi + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \varphi$$

czyli

$$\nabla^2 \varphi - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \varphi = -\frac{\rho}{\epsilon_0}$$

Równania, które uzyskaliśmy są odpowiednikami równań Poissona dla pól stałych i noszą nazwę równań d'Alamberta. Zapisujemy je zwykle w uproszczonej formie wprowadzając nowy operator różniczkowy, nazywany d'alambercjanem:

$$\square = \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}$$

Korzystając z jego definicji zapisujemy równania d'Alamberta w postaci:

$$\square \vec{A} = -\mu_0 \vec{j}$$

$$\square \varphi = -\frac{\rho}{\epsilon_0}$$

Zauważmy, że w przypadku statyki (pola nie zależne od czasu) równania te przechodzą w dyskutowane wcześniej równania Poissona. Tym razem jednak skoncentrujemy uwagę na rozwiązaniach dynamicznych (nie statycznych) czyli zależnych od czasu.

Równanie fali elektromagnetycznej w próżni.

Pole elektromagnetyczne może istnieć i rozprzestrzeniać się również w przypadku gdy w układzie nie występują ładunki i prądy. Wtedy równania na potencjały elektromagnetyczne redukują się do:

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{A} = 0$$

$$\nabla^2 \phi - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \phi = 0$$

Korzystając ze związków pomiędzy wektorami $\vec{E}$ i $\vec{B}$ a potencjałem wektorowym i skalarnym uzyskujemy dla nich również równania d'Alamberta:

$$\nabla^2 \vec{B} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{B} = 0$$

$$\nabla^2 \vec{E} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E} = 0$$

Fala płaska.

Jednym z możliwych rozwiązań równania d'Alamberta jest "fala płaska".

$$\vec{A}(\vec{r}, t) = \vec{A}_0 \sin(\vec{k} \cdot \vec{r} - \omega t)$$

jest rozwiązaniem równania D'Alamberta jeżeli

$$k^2 = \frac{\omega^2}{c^2} \text{ czyli } \omega = kc$$

Zauważmy, że punkty w przestrzeni odpowiadające stałej fazie tego rozwiązania

$$\vec{k} \cdot \vec{r} - \omega t = \text{const}$$

poruszają się z prędkością, która spełnia związek

$$\vec{k} d\vec{r} = \omega dt \text{ czyli } \vec{k} \frac{d\vec{r}}{dt} = \omega = kc$$

czyli prędkość rozprzestrzeniania się fazy (prędkość fazowa jest równa c , a kierunek jest zgodny z kierunkiem wektora $\vec{k}$, który nazywamy wektorem falowym.

Korzystając z wyrażenia na potencjał wektorowy możemy wyliczyć odpowiadający cechowaniu Lorentza potencjał skalarny:

$$\dot{\phi} = -c^2 \text{div} \vec{A} = c^2 \vec{k} \cdot \vec{A}_0 \cos(\vec{k} \cdot \vec{r} - \omega t)$$

$$\phi = \frac{c^2}{\omega} \vec{k} \cdot \vec{A}_0 \sin(\vec{k} \cdot \vec{r} - \omega t) = \frac{c}{k} \vec{k} \cdot \vec{A}_0 \sin(\vec{k} \cdot \vec{r} - \omega t)$$

Korzystając z wyrażen na potencjały możemy wyliczyć wektory pola magnetycznego i elektrycznego:

$$\vec{B} = \nabla \times \vec{A} = \vec{k} \times \vec{A}_0 \cos(\vec{k} \cdot \vec{r} - \omega t)$$

$$\vec{E} = -\nabla \phi - \frac{\partial}{\partial t} \vec{A} = -\frac{c}{k} (\vec{k} \cdot \vec{A}_0) \vec{k} \cos(\vec{k} \cdot \vec{r} - \omega t) + \omega \vec{A}_0 \cos(\vec{k} \cdot \vec{r} - \omega t) =$$

$$= \left(-\frac{c \vec{k} \cdot \vec{A}_0}{k} \vec{k} + c k \vec{A}_0 \right) \cos(\vec{k} \cdot \vec{r} - \omega t)$$

Widzimy, że wektory $\vec{k}$ i $\vec{B}$ są do siebie prostopadłe. Możemy przekonać się, że również wektor $\vec{E}$ jest prostopadły do $\vec{k}$:

$$\vec{k} \cdot \left(-\frac{c \vec{k} \cdot \vec{A}_0}{k} \vec{k} + c k \vec{A}_0 \right) = -\frac{c \vec{k} \cdot \vec{A}_0}{k} k^2 + c k \vec{A}_0 \cdot \vec{k} = 0$$

Podobnie pokazujemy prostopadłość wektorów $\vec{B}$ i $\vec{E}$.

$$\vec{k} \times \vec{A}_0 \cdot \left(-\frac{c\vec{k} \cdot \vec{A}_0}{k} \vec{k} + c\vec{A}_0 \right) = 0$$

Wektory $\vec{k}$, $\vec{B}$ i $\vec{E}$ stanowią trójkę wzajemnie prostopadłych wektorów. Zatem pole elektryczne i magnetyczne są do siebie prostopadłe i prostopadłe do kierunku prędkości rozchodzenia się fazy. Fala elektromagnetyczna jest falą poprzeczną.

Podsumowanie XIII.

- Przeanalizowaliśmy zależności pomiędzy E i D oraz H i B.
- Przedyskutowaliśmy zjawisko polaryzacji dielektryka.
- Przedyskutowaliśmy pola elektryczne i magnetyczne na granicy ośrodków.
- Rozważyliśmy zmienne w czasie pola pochodzące od zmiennych w czasie źródeł.
- Potencjały pól muszą spełniać równania d'Alamberta.
- Przedyskutowaliśmy rozwiązanie równania d'Alamberta w próżni odpowiadające fali płaskiej.
- Faza fali elektromagnetycznej rozchodzi się równoległe do kierunku wektora falowego z prędkością równą prędkości światła.
- Fala elektromagnetyczna jest falą poprzeczną, trzy wektory k , B , E są do siebie wzajemnie prostopadłe.

XIV.

Elektrodynamika klasyczna w nanostrukturach półprzewodnikowych.

Podsumowanie elektrodynamiki.

Z zasady najmniejszego działania dla pól elektrycznego i magnetycznego w obecności rozkładu ładunków i prądów wyprowadziliśmy równania Maxwella. Równania te uzupełnione o wzór na siłę Lorentza (wynikający z tego samego formalizmu), oraz równanie ciągłości stanowią kompletny układ równań elektrodynamiki, z których można uzyskać wszystkie inne prawa. Jednocześnie ich rozwiązania są spójne z mechaniką relatywistyczną.