

Wykłady z relatywistycznej mechaniki
kwantowej (wersja wstępna)

Kacper Zalewski

27 października 2019

Spis treści

1	Szczególna teoria względności	3
1.1	Układy inercjalne i zasada względności	3
1.2	Geometria Minkowskiego	4
1.3	Klasyfikacja Transformacji Lorentza	6
1.4	Właściwe transformacje Lorentza	8
1.5	Przyszłość, przeszłość, przyczynowość	11
1.6	Przykłady zastosowania właściwych transformacji Lorentza	13
1.7	Czterowektory i skalary	14
1.8	Przykłady czterowektorów i skalarów	17
1.9	Przykłady zastosowania czterowektorów i skalarów	19
1.10	Tensory	20
1.11	Uzupełnienie: Liniowość transformacji Lorentza	23
2	Cząstka w polu elektromagnetycznym	25
2.1	Przypomnienie wiadomości z klasycznej elektrodynamiki	25
2.2	Efekt Bohma-Aharonova	27
2.3	Symetria względem cechowania w mechanice kwantowej	29
2.4	Równanie Schrödingera dla cząstki oddziałującej z polem elektromagnetycznym	32
3	Równanie Kleina-Gordona	37
3.1	Równanie Kleina-Gordona i jego rozwiązania dla cząstki swobodnej	37
3.2	Prąd zachowywany	42
3.3	Symetrie równania Kleina-Gordona	43
4	Spinory w szczególnej teorii względności	49
4.1	Grupa $SL(2, C)$, spinory z górnymi wskaźnikami	49
4.2	Spinory z dolnymi wskaźnikami	51
4.3	Spinory ze wskaźnikami kropkowanymi	53
4.4	Tensory spinorowe	55
4.5	Związek tensorów spinorowych z czterowektorami	56
4.6	Właściwe transformacje Lorentza spinorów	59
4.7	Inwersja przestrzenna	61
4.8	Sprzężenie spinora i operacja CP	63
4.9	Odwroćenie czasu i operacja CPT	64
4.10	Bispinory	66
4.11	Chiralność	68
4.12	Kowarianty dwuliniowe	69

4.13	Własności macierzy γ^μ	73
5	Równania relatywistyczne dla cząstek o spinie $S = \frac{1}{2}$	75
5.1	Równania Weyla	75
5.2	Symetrie równań Weyla	77
5.3	Równanie Diraca	79
5.4	Symetrie równania Diraca	82
5.5	Cząstka swobodna – reprezentacja spinorowa	85
5.6	Cząstka swobodna – reprezentacja Diraca	88
5.7	Poprawki relatywistyczne do opisu cząstki w polu elektromagnetycznym: pierwsze przybliżenie	92
5.8	Poprawki relatywistyczne do opisu cząstki w polu elektromagnetycznym: drugie przybliżenie	94
5.9	Paradoks Kleina	98
5.10	Promieniowanie Hawkinga	101
5.11	Cząstki Majorany	102

Rozdział 1

Szczególna teoria względności

1.1 Układy inercjalne i zasada względności

Szczególna teoria względności określa, jak mają się do siebie obserwacje robione przez obserwatorów spoczywających w różnych inercjalnych układach odniesienia. Zaczynamy więc od wprowadzenia układów inercjalnych. Układy inercjalne, które będziemy rozważać, są układami odniesienia w czterowymiarowej czasoprzestrzeni. Punkt w takiej czasoprzestrzeni jest określony przez podanie czterech liczb (t, \mathbf{x}) i nazywa się **zdarzeniem** (ang. event) lub **punktochwilą**. Można sobie wyobrazić, że w każdym punkcie \mathbf{x} zwykłej przestrzeni znajduje się zegar odmierzający czas t . Zdarzenie (t, \mathbf{x}) następuje w punkcie \mathbf{x} , kiedy znajdujący się tam zegar wskazuje czas t . Wszystkie te zegary chodzą tak samo w tym sensie, że dwa zegary umieszczone w tym samym punkcie \mathbf{x} i zsynchronizowane w jakiejś chwili, na przykład $t = 0$, tak że pokazują ten sam czas, pozostają zsynchronizowane dla wszystkich $t > 0$. Mówimy, że dwa zegary w jednym punkcie można zsynchronizować trwale. Synchronizacja zegarów znajdujących się w różnych punktach \mathbf{x} jest bardziej skomplikowana, ale wiadomo jak to robić i nie ma tu istotnych trudności, z tym że taka synchronizacja na ogół nie okazuje się trwała.

Układ odniesienia nazywamy **układem inercjalnym**, jeśli spełnione są następujące dwa warunki [2].

- Da się trwale zsynchronizować zegary we wszystkich punktach \mathbf{x} .
- Dla danej chwili t (według tych zsynchronizowanych zegarów) geometria w przestrzeni \mathbf{x} jest euklidesowa, przy czym odległości między punktami \mathbf{x} nie zależą od wyboru chwili t .

Czasem, zamiast o układzie inercjalnym, mówi się o obserwatorze inercjalnym, lub krócej o obserwatorze, rozumiejąc przez to obserwatora nieruchomego w danym układzie. Dalej będziemy często pisać układ zamiast układ inercjalny. Ponieważ układy, z którymi będziemy mieli do czynienia, będą prawie zawsze inercjalne, ten skrót nie prowadzi do nieporozumień. Można pokazać, że układ poruszający się ruchem przyspieszonym względem układu inercjalnego nie może być układem inercjalnym oraz, że układ w którym działa pole grawitacyjne nie może być układem inercjalnym. Układ inercjalny jest więc pewną idealizacją spotykanych w rzeczywistości układów.

Ważną dodatkową informację o układach inercjalnych można uzyskać w oparciu o **zasadę względności**, którą można sformułować następująco:

- Prawa fizyki są identyczne we wszystkich układach poruszających się względem siebie ruchem jednostajnym prostoliniowym.

Ta zasada nie obejmuje układów otrzymanych przez odbicia przestrzenne lub czasowe, które omówimy w paragrafie 1.3. Z zasady względności wynika natychmiast, że każdy układ poruszający się ruchem jednostajnym prostoliniowym w stosunku do jakiegoś układu inercjalnego też jest układem inercjalnym.

Zasada względności jest jednym z dwu podstawowych założeń, na których opiera się szczególna teoria względności. Skromne początki tej zasady są pięknie opisane w podręczniku [3], gdzie można znaleźć obszerny cytat z pracy Galileusza, od której się wszystko zaczęło, i informacje o pierwszych testach doświadczalnych. Podane tu sformułowanie zasady względności nie stosuje się, jeśli trzeba uwzględnić pola grawitacyjne [2]. Einstein później znalazł uogólnienie, które stosuje się i w takich przypadkach, ale to już należy do wykładu ogólnej teorii względności.

Ćwiczenie Wnętrze sztucznego satelity lecącego z wyłączonymi motorami jest w dobrym przybliżeniu układem inercjalnym. Przedyskutować jakościowo ruch względny dwu sztucznych satelitów poruszających się po różnych orbitach w wywnioskować stąd, że pole grawitacyjne psuje zasadę względności.



1.2 Geometria Minkowskiego

Jako drugie, obok zasady względności, założenie przyjmujemy, że czasoprzestrzeń jest **przestrzenią Minkowskiego** (Minkowski 1908). Opartą o aksjomaty teorii przestrzeni Minkowskiego, ze szczególnym podkreśleniem jej podobieństwa do teorii przestrzeni euklidesowej, można znaleźć w skrypcie [4]. Pojęciem pierwotnym są **wektory**, które będziemy zapisywać zwykłym drukiem, rezerwując druk pogrubiony dla wektorów bez składowej czasowej. Weźmy przykładowo wektor Δx . Symbol Δ jest wprowadzony, żeby przypomnieć, że taki wektor nie musi być zaczepiony w początku układu, a więc może być rozumiany jako różnica dwu wektorów wodzących. **Kiedy wektor jest zaczepiony w początku układu, można go oznaczać x** . Podobnie jak w geometrii euklidesowej, wektor może być określony przez podanie jego składowych, ale to wymaga uprzedniego wyboru układu odniesienia. Stosuje się zapis:

$$\Delta x = \{\Delta x^0, \Delta x^1, \Delta x^2, \Delta x^3\} = \{\Delta x^\mu\}, \quad (1.1)$$

gdzie wskaźniki 0, 1, 2, 3 oznaczają kolejno składowe: czasową, x -ową, y -ową i z -ową. Z kontekstu wnioskuje się, czy w danym zdaniu lub wzorze Δx oznacza niezależny od układu odniesienia wektor, czy zbiór zależnych od układu odniesienia składowych, czy x -ową składową wektora. Przyjmujemy konwencję, że wskaźniki łacińskie przebiegają wartości (1, 2, 3), a wskaźniki greckie wartości (0, 1, 2, 3). Zauważmy, że translacja zmienia punkt zaczepienia wektora, ale nie zmienia jego składowych. Przyjmujemy **układ jednostek**, w którym prędkość światła (w próżni) $c = 1$. Tak więc wszystkie cztery składowe wektora mają ten sam wymiar.

W geometrii euklidesowej wyróżnione są przekształcenia, które nie zmieniają kwadratów długości wektorów. Te przekształcenia, zwane **euklidesowymi**, stanowią grupę i obejmują translacje, obroty i inwersję. Podobnie w geometrii Minkowskiego wyróżniona jest grupa przekształceń, które nie zmieniają formy kwadratowej

interw

$$\Delta s^2 = (\Delta t)^2 - (\Delta \mathbf{x})^2 = g_{\mu\nu} \Delta x^\mu \Delta x^\nu, \quad (1.2)$$

gdzie

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (1.3)$$

Te wzory wymagają komentarzy. W ostatnim wyrażeniu we wzorze (1.2) zastosowana została **konwencja Einsteina**. Według tej konwencji, jeżeli jakiś wskaźnik grecki powtarza się dwa razy, raz jako górny, raz jako dolny, to znaczy że sumuje się po nim od 0 do 3. We wzorze mamy więc sumę 16 członów, odpowiadającą dwu powtórzonym greckim wskaźnikom. Macierz $g_{\mu\nu}$ jest znana jako **tensor metryczny**. Wybraliśmy konwencję, według której znaki elementów diagonalnych są $+- --$. Można wybrać jako tensor metryczny macierz $-g_{\mu\nu}$, co odpowiada konwencji $-++++$. Wśród fizyków zajmujących się fizyką w czterowymiarowej czasoprzestrzeni konwencja $+- --$ wyraźnie wypiera popularniejszą kiedyś konwencję $-++++$, ale w teoriach więcej-wymiarowych konwencja $-++++$ i jej uogólnienia na więcej wymiarów są dogodniejsze¹. W każdym razie, czytając pracę bezpiecznie jest sprawdzić, której konwencji autor używa. Wyrażenie Δs^2 jest znane jako **interwał**. Interwał może być dodatni, ujemny lub zerowy. Trajektorię cząstki w czasoprzestrzeni nazywamy **linią świata** tej cząstki. Dla dowolnie wybranych dwu punktochwili linii świata, interwał Δs^2 jest dodatni, jeśli cząstka porusza się z prędkością (średnio) mniejszą od prędkości światła, ujemny jeśli ruch jest z prędkością większą od prędkości światła i zerowy jeśli prędkość jest równa prędkości światła.

Ćwiczenie Udowodnić stwierdzenia zawarte w ostatnim zdaniu.

•••••

Grupę przekształceń, które nie zmieniają interwałów, będziemy nazywali rozszerzoną grupą Poincarégo. Podgrupa pozostała po usunięciu z tej grupy translacji nazywa się **grupą Lorentza**. W dalszym ciągu będziemy się zajmowali głównie grupą Lorentza. Transformacje Lorentza, z definicji, przekształcają **początek układu** w początek układu. Stwierdzenie, że czasoprzestrzeń jest przestrzenią Minkowskiego, interpretujemy jako założenie, że przejścia między układami inercyjnymi są transformacjami z rozszerzonej grupy Poincarégo, to znaczy że interwał nie zależy od (inercyjnego) układu odniesienia. Udowodnimy teraz ważne twierdzenie fizyczne.

Wybieramy wektor Δx , którego początek i koniec znajdują się w dwu różnych punktochwilach linii świata cząstki poruszającej się ruchem prostoliniowym z prędkością światła $c = 1$. W takim razie

¹Dziękuję prof. K. Meissnerowi za informacje na ten temat

$$(\Delta t)^2 = (\Delta \mathbf{x})^2, \quad (1.4)$$

co oznacza, że

$$\Delta s^2 = 0. \quad (1.5)$$

Ponieważ interwał nie zależy od układu odniesienia, prędkość tej cząstki jest równa prędkości światła w każdym układzie odniesienia. To jest zaskakujący wynik. Cząstki poruszającej się z prędkością światła, w szczególności światła, nie da się dogonić. Nie ma układu, w którym by taka cząstka spoczywała. Einstein przyjął równoważny wynik jako założenie² i uzupełniwszy je zasadą względności wyprowadził całą szczególną teorię względności. Zamiana założenia, że prędkość światła jest taka sama w każdym układzie, na założenie, że czasoprzestrzeń jest przestrzenią Minkowskiego jest typowa dla ewolucji fizyki teoretycznej. Przechodzi się od założeń, które mają prosty sens intuicyjny czy bezpośredni dowód doświadczalny, do założeń, które stanowią dogodny punkt wyjścia do zastosowań. Na przykład, Newton wyprowadził swoje prawo grawitacji z praw Keplera, wczesne sformułowania drugiej zasady termodynamiki orzekały, że nie da się zbudować maszyn pewnych typów.

Warunkiem wystarczającym, żeby przekształcenie należało do rozszerzonej grupy Poincarégo jest, żeby nie zmieniało ani $(\Delta t)^2$, ani $(\Delta \mathbf{x})^2$. Stąd wszystkie przekształcenia euklidesowe, należą do tej grupy. Zauważmy, że przekształcenia euklidesowe nie zmieniają prędkości układu. Pozostają więc do zbadania przejścia między układami inercjalnymi poruszającymi się względem siebie. Takie przekształcenia są nazywane **pchnięciami** (ang. boost).

1.3 Klasyfikacja Transformacji Lorentza

Oznaczmy $\{x^\mu\}$ współrzędne w układzie U jakiejś punktochwili i $\{x'^\mu\}$ współrzędne tej samej punktochwili w układzie U' . Szukamy transformacji

$$x'^\mu = x'^\mu(x^0, x^1, x^2, x^3), \quad \mu = 0, 1, 2, 3 \quad (1.6)$$

nie zmieniających interwałów, a więc w szczególności takich, że

$$g_{\alpha\beta} dx^\alpha dx^\beta = g_{\mu\nu} dx'^\mu dx'^\nu = g_{\mu\nu} \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} dx^\alpha dx^\beta. \quad (1.7)$$

Zaczynamy od ogólnej klasyfikacji. Porównując współczynniki przy $dx^\alpha dx^\beta$ w pierwszym i trzecim wyrażeniu otrzymujemy

$$\frac{\partial x'^\mu}{\partial x^\alpha} g_{\mu\nu} \frac{\partial x'^\nu}{\partial x^\beta} = g_{\alpha\beta}. \quad (1.8)$$

Zauważmy, że lewa strona tej równości jest iloczynem trzech macierzy, przy czym pierwsza i trzecia są względem siebie transponowane³. Biorąc wyznac-

²W swojej pracy z roku 1905 Einstein założył, że prędkość światła nie zależy od prędkości jego źródła. To znaczy, na mocy zasady względności, że jest taka sama w układzie, w którym źródło spoczywa, i w układzie, w którym źródło porusza się z dowolnie ustaloną prędkością $|\mathbf{v}| < 1$.

³Transpozycja jest potrzebna, bo mnożąc macierze sumuje się po przyległych wskaźnikach. Jeśli w pochodnej wskaźnik górny uznamy za pierwszy, to drugi iloczyn po lewej stronie jest iloczynem macierzy, natomiast pierwszy wymaga transpozycji macierzy stojącej po lewej stronie

niki z obu stron równości, i wykorzystując twierdzenia, że wyznacznik z iloczynu macierzy jest równy iloczynowi wyznaczników tych macierzy oraz że transpozycja macierzy nie zmienia jej wyznacznika, mamy wobec $\text{Det } g_{\mu\nu} = -1$, lordet

$$\left(\text{Det} \frac{\partial x'^{\nu}}{\partial x^{\beta}} \right)^2 = 1. \quad (1.9)$$

Są więc dwie możliwości: wyznacznik w powyższym wzorze może być równy $+1$ lub -1 . Kładąc we wzorze (1.8) $\alpha = \beta = 0$ otrzymujemy:

$$\left(\frac{\partial x'^0}{\partial x^0} \right)^2 = 1 + \left(\frac{\partial x'^1}{\partial x^0} \right)^2 + \left(\frac{\partial x'^2}{\partial x^0} \right)^2 + \left(\frac{\partial x'^3}{\partial x^0} \right)^2 \geq 1. \quad (1.10)$$

I znowuż powstają dwie możliwości: $\frac{\partial x'^0}{\partial x^0} \geq 1$, lub $\frac{\partial x'^0}{\partial x^0} \leq -1$. Możemy więc wyróżnić następujące cztery rodzaje transformacji Lorentza [4].

- Transformacje właściwe⁴, dla których ortoch

$$\text{Det} \frac{\partial x'^{\nu}}{\partial x^{\mu}} = 1; \quad \frac{\partial x'^0}{\partial x^0} \geq 1. \quad (1.11)$$

Tę grupę transformacji będziemy oznaczali **symbolem** $L_{\uparrow+}$.

- Odbicia czasowe, dla których

$$\text{Det} \frac{\partial x'^{\nu}}{\partial x^{\mu}} = -1; \quad \frac{\partial x'^0}{\partial x^0} \leq -1, \quad (1.12)$$

- Odbicia przestrzenne, dla których

$$\text{Det} \frac{\partial x'^{\nu}}{\partial x^{\mu}} = -1; \quad \frac{\partial x'^0}{\partial x^0} \geq 1. \quad (1.13)$$

- Odbicia zupełne, dla których

$$\text{Det} \frac{\partial x'^{\nu}}{\partial x^{\mu}} = 1; \quad \frac{\partial x'^0}{\partial x^0} \leq -1. \quad (1.14)$$

Tylko **właściwe transformacje Lorentza** przekształcają układy inercjalne w układy inercjalne w rzeczywistym świecie. Dlatego, głównie te będziemy dyskutowali i, tam gdzie to nie prowadzi do nieporozumienia, będziemy je nazywali po prostu transformacjami Lorentza. Pozostałe transformacje, które obejmujemy łączną nazwą niewłaściwych transformacji Lorentza lub odbić, przekształcają rzeczywiste układy inercjalne w układy nierzeczywiste, na przykład dla odbić przestrzennych, w układy po drugiej stronie lustra. W tych innych światach inne prawa rządzą słabymi oddziaływaniami, nie są to więc te układy inercjalne, dla których jest spełniona zasada względności. Mimo to, jak dalej zobaczymy, badanie niewłaściwych transformacji Lorentza jest użyteczne. Można je łatwo otrzymać z właściwych, dokonując dodatkowo: dla odbić czasowych transformacji $\Delta t \rightarrow -\Delta t$, dla odbić przestrzennych transformacji $\Delta \mathbf{x} \rightarrow -\Delta \mathbf{x}$ i dla odbić

⁴Zwykle dodaje się ortochroniczne, lub z zachowaniem kierunku czasu, ale pozostaniemy przy krótszej nazwie

zupełnych transformacji $\Delta t \rightarrow -\Delta t$ i $\Delta \mathbf{x} \rightarrow -\Delta \mathbf{x}$. Zauważmy jeszcze, że niewłaściwe transformacje Lorentza, w odróżnieniu od właściwych, nie zawierają jedynki (identyczności), nie stanowią więc podgrup pełnej grupy Lorentza.

Najmniejszą grupę zawierającą translacje i transformacje $L_{\uparrow+}$ nazywamy **grupą Poincarégo**. Założenie szczególnej teorii względności możemy więc sformułować następująco: **Transformacja z układu inercjalnego prowadzi do układu inercjalnego równoważnego w sensie zasady względności, wtedy i tylko wtedy, kiedy należy do grupy Poincarégo.**

Przedyskutujemy teraz bardziej szczegółowo właściwe transformacje Lorentza.

1.4 Właściwe transformacje Lorentza

Zauważmy, że jeśli poddamy tej samej transformacji czterowektor i wektory bazowe układu odniesienia, to składowe nowego czterowektora względem nowej bazy będą identyczne ze składowymi starego wektora względem starej bazy. Jeśli chcemy otrzymać jakąś zmianę, to możemy albo przetransformować czterowektor pozostawiając wektory bazowe, czyli układ odniesienia, bez zmian, albo przetransformować układ odniesienia nie zmieniając czterowektora.

Ćwiczenie Sprawdzić te stwierdzenia dla obrotów wokół osi z .

•••••

Dotychczas stosowaliśmy tak zwaną **interpretację bierną** transformacji Lorentza: transformacja przeprowadzała składowe wektora w układzie U w składowe tego samego wektora w układzie U' , a więc zmieniały się tylko wektory bazowe. Przy **interpretacji czynnej**, transformacja przeprowadza składowe wektora z układu U' w składowe tego samego wektora, w układzie U . Jest to więc transformacja odwrotna w stosunku do transformacji przy interpretacji biernej. Podręcznikowe macierze $D_{m'm}^j(\alpha, \beta, \gamma)$, reprezentujące obroty, są podawane dla interpretacji czynnej, dlatego począwszy od tego paragrafu tę interpretację będziemy stosować.

Określenia reprezentacji, czynna i bierna, wiążą się z następującym rozumowaniem. Niech transformacja interpretowana biernie przeprowadza wektor $\Delta x'$ w wektor Δx . Przez wektor rozumiemy tu zbiór jego składowych, bo sam wektor pozostaje bez zmian, jest bierny. Obserwator natomiast zaczyna używać układu U' zamiast układu U , co zmienia wektory bazowe i, co za tym idzie, współrzędne wektora. Opiszemy teraz tę samą transformację w interpretacji czynnej. W układzie U , oprócz wektora $\Delta x'$, definiujemy nowy wektor, który ma takie same składowe w układzie U , jak wektor Δx w układzie U' . Ponieważ transformacja w interpretacji czynnej jest odwrotna do transformacji w interpretacji biernej, przy interpretacji czynnej składowe wektora Δx przechdzą w składowe wektora $\Delta x'$. Oba wektory są zdefiniowane w tym samym układzie, a więc wektory bazowe nie ulegają zmianie i to wektor zostaje zmodyfikowany. W tym sensie bierze czynny udział w transformacji.

Ćwiczenie Narysować dowolny wektor $\mathbf{x}' = \{x', y'\}$ zaczepiony w początku układu U na płaszczyźnie. Dorysować układ U' , którego osie powstają z osi układu U przez obrót o kąt α wokół osi z . Składowe wektora \mathbf{x}' w układzie U' oznaczyć $\{x, y\}$. Pokazać, że transformacja $\{x', y'\} \rightarrow \{x, y\}$ jest obrotem o kąt α w interpretacji biernej. Narysować układ U z wektorami $\{x, y\}$ i $\{x', y'\}$.

Pokazać, że transformacja $\{x, y\} \rightarrow \{x', y'\}$ jest obrotem o kąt α w interpretacji czynnej.

•••••

Ćwiczenie Sprawdzić, że wyniki uzyskane w poprzednich dwu paragrafach nie ulegają zmianie przy zamianie interpretacji biernej na czynną.

•••••

Każdą właściwą transformację Lorentza $U' \rightarrow U$ można sprowadzić do ciągu trzech transformacji: obrotu euklidesowego (to znaczy w trójwymiarowej przestrzeni euklidesowej), po którym oś z jest ustawiona równoległe do pchnięcia, pchnięcia wzdłuż osi z i obrotu euklidesowego ustawiającego osie przestrzenne w kierunkach x', y', z' .

Ćwiczenie Pokazać, że ustawienie osi x', y', z' określa jednoznacznie położenie osi t' .

•••••

Oczywiście, jeśli układy U i U' poruszają się z tą samą prędkością, to cała transformacja sprowadza się do obrotu i pierwsze dwie operacje są zbędne. Ponieważ przyjmujemy, że teoria obrotów jest znana, wystarczy przedyskutować pchnięcia wzdłuż osi z . Wszystkie wprowadzane w tym paragrafie prędkości układów będą równoległe, lub antyrównoległe, do osi z .

Nie zmniejszając ogólności możemy przyjąć, że pchnięcie wzdłuż osi z jest **bezobrotowe**, to znaczy takie, że

bezobr

$$\Delta x = \Delta x' \quad \text{i} \quad \Delta y = \Delta y'. \quad (1.15)$$

Warunek zachowania interwału przybiera więc postać

intran

$$(\Delta t)^2 - (\Delta z)^2 = (\Delta t')^2 - (\Delta z')^2. \quad (1.16)$$

Zakładając, że relacja między składowymi Δx i $\Delta x'$ jest liniowa⁵ i spełnia warunki (1.11) można ją zapisać w postaci

loreta

$$\Delta t' = \Delta t \cosh \eta + \Delta z \sinh \eta, \quad \Delta z' = \Delta t \sinh \eta + \Delta z \cosh \eta. \quad (1.17)$$

Wprowadzony tu parametr η jest nazywany **pośpiesznością** (ang. rapidity) i ma prostą interpretację fizyczną. Wybierzmy jako wektor Δx odcinek linii świata cząstki spoczywającej w układzie U' , a zatem poruszającej się w układzie U z prędkością v równą prędkości układu U' względem układu U . To daje $\Delta z = 0$. Podstawiając do wzorów (1.17) i dzieląc stronami równanie drugie przez pierwsze otrzymujemy

defrap

$$\tanh \eta = \frac{\Delta z'}{\Delta t'} = v, \quad (1.18)$$

Mamy więc, wzajemnie jednoznaczność między pośpiesznością i prędkością układu U' , a więc i cząstki, względem układu U . Przy okazji udowodniliśmy, że dla danego pchnięcia, określonego przez pośpieszność η , układ

⁵To założenie można udowodnić. Dowód podany jest w uzupełnieniu na końcu tego rozdziału.

U' porusza się względem układu U ruchem jednostajnym tak, jak to powinno być dla dwu układów inercjalnych.

Ćwiczenie Napisać, dla pchnięć bezobrotowych, ogólną relację liniową między składowymi Δx i $\Delta x'$, a następnie wykazać, korzystając z tożsamości (1.16) i warunków (1.11), że można ją zapisać w postaci (1.17).

•••••

lorenz

Wyrażając kosinus hiperboliczny przez tangens hiperboliczny i korzystając ze wzoru (1.18), możemy przepisać wzory (1.17) w postaci:

$$\begin{aligned}\Delta t' &= \gamma(\Delta t + v\Delta z), \\ \Delta z' &= \gamma(\Delta z + v\Delta t),\end{aligned}\tag{1.19}$$

gamma

gdzie

$$\gamma = \frac{1}{\sqrt{1-v^2}}.\tag{1.20}$$

Ta funkcja prędkości jest znana jako **czynnik Lorentza**.

Transformacją odwrotną do pchnięcia bezobrotowego z prędkością v jest pchnięcie bezobrotowe z prędkością $-v$, a zatem żeby przejść do interpretacji biernej wystarczy zmienić znak prędkości. Wzory (1.19) ze zmienionym znakiem prędkości spotyka się wielu w podręcznikach.

Ćwiczenie W układzie U , wzdłuż osi z' , jedzie samochód z prędkością \mathbf{v} . Z samochodem wiążemy układ U' tak, że dla środka samochodu $z \equiv 0$. Sprawdzić w przybliżeniu nierelatywistycznym wzory (1.19).

•••••

Ćwiczenie Pokazać, że w układzie U oś t tworzy kąt $\theta = \arctan v$ z osią t' i kąt $\frac{\pi}{2} - 2\theta$ z osią z .

•••••

Wzór (1.20) traci sens dla $v > 1$, więc cząstka punktowa nie może się poruszać z prędkością większą od prędkości światła⁶, i co za tym idzie nie może być użyta do przekazania informacji z prędkością większą od prędkości światła. Popularna hipoteza, że w żaden sposób nie da się przekazać informacji z prędkością większą niż prędkość światła występuje już u Poincarégo [3] i dla realistycznych układów może być prawdziwa, ale nie należy do standardowej szczególnej teorii względności. Einstein przyjął tylko założenie [3], że prędkość światła nie zależy od prędkości jego źródła. Przy tym założeniu można udowodnić powyższą hipotezę dla rozpatrywanej powyżej cząstki punktowej, czy dla elektrodynamiki, ale także wskazać kontrprzykłady. Na przykład, w relatywistycznej hydrodynamice można wstawić dowolną prędkość dźwięku, także większą od prędkości światła⁷. Ważne jest tu, że chodzi o przesłanie informacji, bo ruch z prędkością większą od prędkości światła jest możliwy. Można się o tym przekonać, choćby z

⁶Rozważa się niekiedy tak zwane tachyony – cząstki punktowe, które mogą mieć różne prędkości, ale zawsze większe od prędkości światła. Istnienie tachyonów nie jest wykluczone przez podane tu rozumowanie

⁷Dziękuję prof. Andrzejowi Staruszkiewiczowi za dyskusję i informacje na ten temat.

następującego doświadczenia myślowego. Wskaźnik laserowy osadzony w punkcie $\mathbf{x} = \mathbf{0}$ obraca się z prędkością kątową ω i rzuca plamkę świetlną na sferę $\mathbf{x}^2 = R^2$. Prędkość liniowa tej plamki wynosi ωR , i przy dostatecznie dużym R może przekroczyć prędkość światła. Jednak ta plamka nie przenosi żadnej informacji.

Ćwiczenie Pokazać, że kiedy jeden układ porusza się względem drugiego z prędkością bliską prędkości światła ($\eta \rightarrow \pm\infty$), zachodzi

$$\Delta t' \pm \Delta z' = \frac{1}{2}(\Delta t \pm \Delta z)e^{\pm\eta}, \quad (1.21)$$

gdzie górne znaki są dla $\eta \rightarrow \infty$ a dolne dla $\eta \rightarrow -\infty$.

•••••

Rozważmy teraz problem dodawania prędkości. Na przykład, dróżnik stojący przy torze kolejowym obserwuje podróżnego idącego wzdłuż korytarza pociągu. Prędkość wagonu względem dróżnika wynosi v_1 i pasażera względem wagonu v_2 . Jaka jest prędkość v_{suma} podróżnego względem dróżnika? Zakładamy, że wszystkie prędkości są równoległe do osi z . Potrzebne będą wzory (1.17) i analogiczny wzór

$$\Delta t'' = \Delta t' \cosh \eta' + \Delta z' \sinh \eta', \quad (1.22)$$

gdzie η' jest pospiesznością pasażera względem wagonu. Eliminując $\Delta t'$ i $\Delta z'$ otrzymujemy

$$\Delta t'' = \Delta t \cosh(\eta' + \eta) + \Delta z \sinh(\eta' + \eta), \quad (1.23)$$

gdzie skorzystaliśmy ze wzorów na sinus hiperboliczny i kosinus hiperboliczny sumy. Porównując z pierwszym ze wzorów (1.17), widać że

$$\eta_{suma} = \eta + \eta'. \quad (1.24)$$

A więc **pośpieszności są addytywne**. W tłumaczeniu na prędkości, korzystając ze wzoru na tangens hiperboliczny sumy:

$$v_{suma} = \frac{\tanh \eta + \tanh \eta'}{1 + \tanh \eta \tanh \eta'} = \frac{v + v'}{1 + vv'}. \quad (1.25)$$

Poprawka vv' w mianowniku jest rzędu c^{-2} . Pomijając ją jako małą, odzyskujemy wzór nierelatywistyczny. Zauważmy, że ponieważ dla każdego skończonego η zachodzi $|\tanh \eta| < 1$, zaczynając od cząstki z prędkością $v < 1$ i przyspieszając ją, to znaczy dodając do jej prędkości prędkości **adt**, gdzie **a** jest przyspieszeniem, nigdy nie osiągniemy prędkości $v = 1$.

Ponieważ transformacja Lorentza do każdej pośpieszności dodaje tę samą liczbę, **różnica dwu pośpieszności nie zmienia się przy transformacjach Lorentza** (bezobrotowych, wzdłuż osi z).

1.5 Przyszłość, przeszłość, przyczynowość

Zacniemy od rozważenia sytuacji w czasoprzestrzeni 1 + 1 wymiarowej, czyli z jednym wymiarem przestrzennym, który oznaczymy z . W fizyce nierelatywistycznej, dla punktochwili $\{0, 0\}$, wszystkie punktochwile, które mają współrzędną

t ujemną, należą do przeszłości. Podobnie, wszystkie punktochwile które mają współrzędną t dodatnią, należą do przyszłości. Punktochwile na osi z są równoczesne z punktochwilą $\{0, 0\}$. Ponieważ przy transformacjach Galileusza czas się nie zmienia, to czy jakaś punktochwila x_1 jest wcześniejsza, równoczesna, czy późniejsza, w stosunku do jakiejś innej punktochwili x_2 nie zależy od układu odniesienia. Zobaczmy teraz, że w fizyce relatywistycznej sytuacja jest bardziej skomplikowana.

Rozważmy układ U' poruszający się względem U z prędkością v . Czas w układzie U dany jest wzorem (1.19)

$$t' = \gamma(t + vz).$$

Zgodnie z konwencją wprowadzoną w paragrafie 1.2, pominieliśmy znak Δ dla wektorów zaczepionych w początku układu. Jeśli dla jakiejś punktochwili

$$\Delta s^2 = t^2 - z^2 < 0, \tag{1.26}$$

to można dobrać prędkość $|v| < 1$ tak, że $t' > 0$, lub tak, że $t' < 0$. Dla obserwatora w początku układu U , wybrana punktochwila w pierwszym przypadku należy do przyszłości, a w drugim do przeszłości. Interwały mniejsze od zera nazywamy **interwałami przestrzennymi**. Mamy więc stwierdzenie, że jeżeli interwał między punktochwilami A i B , Δs_{AB}^2 , jest przestrzenny, nie da się w sposób niezależny od wyboru układu odniesienia powiedzieć, która z tych punktochwili jest wcześniejsza. Oczywiście każdy obserwator, w swoim układzie odniesienia, może odpowiedzieć na to pytanie, ale odpowiedzi różnych obserwatorów będą różne.

Ćwiczenie Oznaczając określony w układzie U kąt między osiami z' i z przez θ sprawdzić, że $\tan \theta = v$ i pokazać, że zawsze można dobrać prędkość v tak, żeby oś z przechodziła powyżej (lub poniżej) dowolnie wybranej punktochwili, dla której $z'^2 > t'^2$.

•••••

Interwałami czasowymi nazywamy interwały, dla których

$$\Delta s^2 > 0. \tag{1.27}$$

Jeżeli interwał od punktochwili A do punktochwili B jest czasowy, to wszyscy obserwatorzy są zgodni co do tego, która z tych punktochwili jest wcześniejsza, a która późniejsza. Zdarzenia przyszłe w stosunku do punktochwili $\{0, 0\}$ leżą w trójkącie $t > |z|$. Jeśli wprowadzimy jeszcze jeden wymiar przestrzenny, oś y prostopadłą do osi z i t , ten trójkąt przechodzi w stożek — **stożek przyszłości**. Także dla pełnej, $3 + 1$ wymiarowej czasoprzestrzeni mówi się o stożku przyszłości. Analogicznie zdefiniowany jest **stożek przeszłości**, dla którego $t < -|x|$.

Ważnym pojęciem w fizyce jest **przyczynowość**. Zdarzenie w punktochwili A może być przyczyną zdarzenia w punktochwili B , lub równoważnie zdarzenie w punktochwili B skutkiem zdarzenia w punktochwili A , tylko jeśli wszyscy obserwatorzy są zgodni, że zdarzenie A jest wcześniejsze niż zdarzenie B . Wynika stąd, że jeżeli interwał $\Delta s_{AB}^2 < 0$, to zdarzenie A nie może być ani przyczyną, ani skutkiem zdarzenia B .

1.6 Przykłady zastosowania właściwych transformacji Lorentza

Rozpatrzmy teraz dwa przykłady ilustrujące użyteczność otrzymanych wyników.

Pod wpływem promieni kosmicznych, w górnych warstwach atmosfery, to znaczy na wysokościach rzędu 10km nad poziomem Ziemi, powstają miony. Średni czas życia mionu τ_μ , który można znaleźć w odpowiednich tablicach, wynosi około $2,2 \times 10^{-8}$ sekundy. Rozumując nierelatywistycznie, można by twierdzić, że skoro mion nie może się poruszać z prędkością większą niż prędkość światła, musi się rozpaść przebywszy co najwyżej około $c\tau_\mu \approx 660$ metrów, nie ma więc szans na dotarcie do powierzchni Ziemi. Tymczasem doświadczalnie stwierdza się, że znaczny procent powstałych mionów dochodzi do Ziemi.

Rozumując relatywistycznie, trzeba przede wszystkim zauważyć, że czas życia mionu jest zdefiniowany w jego układzie spoczynkowym U' , a czas potrzebny mu na dolecenie do powierzchni Ziemi w układzie związanym z Ziemią U . W układzie U' mion nie zmienia swojego położenia. Mamy więc $\Delta z \equiv 0$ i ze wzorów (1.19) na transformację Lorentza otrzymujemy w układzie U

dylata

$$\Delta t' = \gamma \Delta t, \quad (1.28)$$

gdzie Δt jest czasem życia mionu w jego układzie spoczynkowym. Czas Δt jest rzędu τ_μ , ponieważ jednak produkowane miony mają wysokie energie i, co za tym idzie, duże współczynniki Lorentza γ , nic dziwnego, że docierają do powierzchni Ziemi. Wzór (1.28) stosuje się nie tylko do mionów. Ogólnie mówimy o relatywistycznej **dylatacji czasu**. W układzie poruszającym się czas płynie wolniej.

Nieco bardziej skomplikowanym przykładem dylatacji czasu jest słynny paradoks bliźniąt. Wystarczy tu rozważyć dwuwymiarową czasoprzestrzeń $\{t, z\}$. W punktochwili O , którą wygodnie jest wybrać jako początek układu, znajdują się dwaj bracia bliźniacy: A i B . B porusza się ze stałą prędkością v względem układu spoczynkowego A . Kiedy na zegarku B upłynie czas τ , prędkość B zmienia się na $-v$, i po dalszym czasie τ bliźniacy spotykają się. Według zegarka B podróż trwała 2τ lat. Ten czas określa, o ile postarzał się B w czasie podróży. Obliczymy teraz czas trwania podróży według zegarka A . Ze wzoru (1.28) na dylatację czasu wynika, że dla A każdy z dwu odcinków podróży B trwał $\gamma\tau$ lat. Cała podróż trwała więc $2\gamma\tau$ lat, i o tyle postarzał się A . Dla $v \approx 1$ jest więc w momencie spotkania znacznie starszy od swojego brata bliźniaka. Zobaczmy jeszcze, jak interpretuje ten wynik B . Stosując wzór (1.28) w układzie spoczynkowym B widać, że dwu odcinkom podróży odpowiadają czasy τ/γ . Tak więc w tych okresach A starzeje się wolniej niż B . Pozostały czas na zegarku A , $2\tau - 2\tau/\gamma$, odpowiada bardzo krótkiemu, według zegarka B , okresowi, kiedy prędkość B zmienia się na przeciwną, czyli kiedy B zmienia swój układ spoczynkowy. Zauważmy jeszcze, że opowieść ilustrująca ten przykład jest mocno wyidealizowana. Zmiana prędkości B w bardzo krótkim czasie wymagałaby ogromnego przyspieszenia, równoważnego włączeniu na ten czas bardzo silnego pola grawitacyjnego, i brat B zostałby zmiażdżony.

Jako kolejny przykład rozważmy dwie rakiety, znajdujące się jedna za drugą, połączone prętem, które według obserwatora na ziemi startują z kosmodromu w tej samej chwili i mają w każdej chwili taką samą prędkość. Przyjmujemy, że

ich prędkości są równoległe do osi z . W każdym układzie, przez długość pręta rozumiemy odległość między punktochwilami na jego końcach, przy założeniu, że te punktochwile mają równe sobie współrzędne czasowe. W tych warunkach, według obserwatora stojącego na ziemi, długość pręta, który łączy rakiety, nie zmienia się. Mimo to, pręt po jakimś czasie zostaje rozerwany. W dyskusji relatywistycznej trzeba rozróżnić l_0 , długość pręta w jego układzie spoczynkowym U' , gdzie stosują się prawa fizyki nierelatywistycznej, i l , długość pręta w układzie U związanym z obserwatorem stojącym na Ziemi. Z założenia, długość l jest stała, ale to ewentualne zmiany długości l_0 mogą doprowadzić do rozerwania pręta.

Rozważmy linie świata końców pręta w układzie U' . Ponieważ w tym układzie pręt spoczywa, te linie są równoległe do osi t . Wybieramy punktochwile x_1 na linii świata początku pręta i punktochwile x_2 na linii świata końca pręta. Dla każdej takiej pary punktochwil

$$\Delta z \equiv z_2 - z_1 = l_0. \quad (1.29)$$

O czasach t_1 i t_2 nic na razie nie zakładamy. Ze wzorów (1.19) otrzymujemy

$$\Delta z' = \gamma(l_0 + v\Delta t); \quad \Delta t' = \gamma(\Delta t + vl_0). \quad (1.30)$$

Długość pręta w układzie U jest równa $\Delta z'$, przy założeniu, że $\Delta t' = 0$. Wyznaczając Δt przy tym warunku z drugiego równania i podstawiając do pierwszego otrzymujemy

$$l = \gamma(l_0 - v^2 l_0) = \frac{l_0}{\gamma}. \quad (1.31)$$

Jeżeli l jest stałe, to l_0 musi rosnać proporcjonalnie do γ , co powoduje w końcu rozerwanie pręta. Wzór (1.31) opisuje tak zwane **skrócenie Lorentza**. Długość poruszającego się obiektu jest mniejsza niż jego długość w układzie spoczynkowym.

Skrócenia Lorentza nie da się zobaczyć. Patrząc na lecący pręt porównujemy punktochwile na końcach pręta, z których światło jednocześnie dociera do naszego oka, a nie te, które są jednoczesne. To prowadzi do zupełnie innych, często zaskakujących wyników. Przystępną dyskusję tych efektów można znaleźć w artykule Weisskopfa [5].

Dobrym ćwiczeniem na skrócenie Lorentza, i nie tylko, jest dyskusja paradoksu stodoły. Tyczka o długości l_1 wpada do stodoły o głębokości $l_2 < l_1$. Jeżeli się cała znajdzie wewnątrz, drzwi stodoły zamykają się. Decyzja obserwatora, czy drzwi się zamknęły, nie może zależeć od układu odniesienia. Tymczasem, można argumentować, że dla obserwatora stojącego przy stodole tyczka, o ile leci dostatecznie szybko, skraca się na tyle, że może cała wejść do stodoły i drzwi się zamykają, a dla obserwatora biegnącego równo z tyczką, to stodoła się skraca i tyczka nie może się w niej zmieścić, więc drzwi się nie zamykają.

1.7 Czterowektory i skalary

W poprzednim paragrafie wyprowadzaliśmy wzory relatywistyczne, używając transformacji Lorentza w czasoprzestrzeni. Wprowadzimy teraz uogólnienie, które pozwala na rozwiązywanie znacznie szerszego zakresu problemów.

W ogólnym przypadku, transformację Lorentza dla składowych wektora Δx możemy zapisać w postaci

lorgor

$$\Delta x'^{\mu} = \Lambda^{\mu}_{\nu} \Delta x^{\nu}, \quad (1.32)$$

gdzie Λ jest odpowiednią macierzą, zależną od kątów obrotów, prędkości pchnięcia i od tego jakich odbić dokonujemy. Ta transformacja jest liniowa, bo jest złożeniem liniowych transformacji: obrotów i pchnięcia bezobrotowego wzdłuż osi z oraz ewentualnie zmiany znaku $\Delta \mathbf{x}$ czy Δt , czy obu. Każdą czwórkę wielkości $\{a^{\mu}\}$, która przy transformacji Lorentza transformuje się według analogicznego wzoru:

czterw

$$a'^{\mu} = \Lambda^{\mu}_{\nu} a^{\nu}, \quad (1.33)$$

nazywamy **czterowektorem**. Czterowektorem jest więc z definicji $\{\Delta x^{\mu}\}$ i co za tym idzie $\{x^{\mu}\}$. Dla każdego innego kandydata, trzeba udowodnić, że transformuje się według wzoru (1.33).

Interwał

$$\Delta s^2 = g_{\mu\nu} \Delta x^{\mu} \Delta x^{\nu} \quad (1.34)$$

nie zmienia się przy transformacjach Lorentza – jest **skalarem**, ściślej mówiąc, skalarem względem transformacji Lorentza. Słowo skalar może być zastąpione słowem **niezmiennik**. Ponieważ wszystkie czterowektory transformują się tak samo, skalarem dla dowolnego czterowektora a jest także

$$a^2 \equiv g_{\mu\nu} a^{\mu} a^{\nu}. \quad (1.35)$$

Ten skalar nazywamy kwadratem czterowektora a . **Iloczyn skalarny**, albo krócej iloczyn, dwu czterowektorów a i b ,

$$ab \equiv g_{\mu\nu} a^{\mu} b^{\nu}, \quad (1.36)$$

też jest skalarem, bo $ab = \frac{1}{2}[(a+b)^2 - a^2 - b^2]$.

Dotychczas rozważaliśmy czterowektory ze wskaźnikiem μ pisanym u góry. Wygodnie jest wprowadzić jeszcze **czterowektory ze wskaźnikiem dolnym**⁸, zdefiniowane wzorem

obnwsk

$$a_{\mu} = g_{\mu\nu} a^{\nu}. \quad (1.37)$$

lub równoważnie

obnws1

$$a_0 = a^0; \quad a_i = -a^i, \quad i = 1, 2, 3. \quad (1.38)$$

Ta operacja nazywa się **obniżaniem wskaźnika**. Powtórzenie jej, jak widać najprościej ze wzorów (1.38), prowadzi z powrotem do wektora z górnymi wskaźnikami. Wzory (1.38) można więc stosować zarówno do podnoszenia jak i do obniżania poziomu wskaźnika. Mamy stąd przepis na **podnoszenie wskaźników**

$$a^{\mu} = \sum_{\nu} g_{\mu\nu} a_{\nu}. \quad (1.39)$$

⁸Spotyka się też inne nazwy: wektory ze wskaźnikiem górnym bywają nazywane kontrawariantnymi, a wektory ze wskaźnikiem dolnym kowariantnymi, lub dualnymi.

Znak sumy jest tu potrzebny, bo oba wskaźniki ν znajdują się na tym samym poziomie, a więc konwencja Einsteina nie stosuje się. Przyjmujemy, że operacje obniżania i podnoszenia wskaźników mogą być tak samo stosowane do wskaźników macierzy g jak do wskaźników czterowektorów. Przedyskutujemy to założenie w paragrafie 1.10. Wynika stąd w szczególności, że

$$g_{\mu\nu} = g^{\mu\nu} \quad (1.40)$$

i poprzedni wzór można zapisać krócej

$$a^\mu = g^{\mu\nu} a_\nu. \quad (1.41)$$

Zauważmy jeszcze, że podnosząc tylko jeden wskaźnik w $g_{\mu\nu}$, lub obniżając go w $g^{\mu\nu}$, otrzymujemy deltę Kroneckera

$$g_{\nu\alpha} g^{\alpha\mu} = \delta_\nu^\mu. \quad (1.42)$$

W równości zawsze można po obu stronach podnieść, lub obniżyć, ten sam wskaźnik. Na przykład równość $a^\mu = b^\mu$ jest równoważna z równością $a_\mu = b_\mu$.

Wzór (1.37) pozwala na następujące zapisanie wzoru na iloczyn skalarny:

$$ab = a_\mu b^\mu = a^\mu b_\mu. \quad (1.43)$$

W podobnych wyrażeniach, sumowanie po wskaźniku czterowektorowym, który występuje raz u góry i raz u dołu, nazywa się **zwięźaniem** i, jak widać, upraszcza własności transformacyjne. Wynik zwięźania nie zależy od tego, który z wskaźników jest u góry, a który u dołu.

Zmieniając odpowiednio poziomy wskaźników, otrzymujemy ze wzoru (1.33)

$$a'_\mu = \Lambda_\mu^\nu a_\nu. \quad (1.44)$$

Macierz Λ_μ^ν jest zdefiniowana przez ten wzór. Stąd

$$a'_\mu b'^\mu = a_\nu \Lambda_\mu^\nu \Lambda^\mu_\alpha b^\alpha = a_\nu b^\nu, \quad (1.45)$$

gdzie ostatnia równość wynika z niezmienniczości iloczynu skalarnego wektorów. Wybierając, w szczególności, wektory a i b tak, żeby wśród 16 iloczynów $a_\rho b^\sigma$ tylko iloczyn $a_\nu b^\nu$ był różny od zera, otrzymujemy

$$\Lambda_\mu^\nu \Lambda^\mu_\alpha = \delta_\alpha^\nu, \quad (1.46)$$

czyli

$$\Lambda_\mu^\nu = (\Lambda^{-1})^\nu_\mu \quad (1.47)$$

i co za tym idzie

$$a'_\mu = a_\nu (\Lambda^{-1})^\nu_\mu. \quad (1.48)$$

Przy tym zapisie niezmienniczość iloczynu skalarnego dwu wektorów jest łatwo widoczna.

Ćwiczenie Uzasadnić to stwierdzenie.



Zauważmy, że własności transformacyjne a_μ zostały tu wyprowadzone z faktu, że $a_\mu b^\mu$ jest skalar. W dowodzie wykorzystaliśmy to, że wszystkie czwórki liczb $\{a_\mu\}$ transformują się tak samo przy transformacjach Lorentza i że da się wybrać czwórki liczb $\{a_\mu\}$ i $\{b^\mu\}$ tak, jak przy wyprowadzaniu wzoru (1.46).

1.8 Przykłady czterowektorów i skalarów

Dla dwu bardzo bliskich sobie punktochwil na linii świata cząstki możemy wyliczyć interwał $ds^2 = dt^2 - dx^2 = dt^2(1 - \mathbf{v}^2)$ i zdefiniować skalar

$$d\tau \equiv \sqrt{ds^2} = \gamma^{-1} dt. \quad (1.49)$$

Z dyskusji w paragrafie 1.5 wynika, że $\mathbf{v}^2 \leq 1$, a zatem zawsze $ds^2 \geq 0$. Skalar τ nazywa się **czasem własnym cząstki** i pozwala zdefiniować dalsze czterowektory.

Użytecznym czterowektorem jest **czteropędność**

$$u^\mu = \frac{dx^\mu}{d\tau} = \{\gamma, \gamma \mathbf{v}\}, \quad (1.50)$$

gdzie pochodne są liczone wzdłuż linii świata cząstki. Zauważmy, że kwapre

$$u^2 = 1. \quad (1.51)$$

Czteropędność jest określona tylko dla cząstek poruszających się z prędkością mniejsza od prędkości światła, bo dla $dt = \pm |d\mathbf{x}|$ mamy $d\tau \equiv 0$ i pochodna po τ traci sens. Do wektora \mathbf{x} wystarczy dołączyć czwartą składową t , żeby zrobić z niego czterowektor. Dla prędkości procedura jest bardziej skomplikowana. Jeszcze bardziej skomplikowany przypadek spotkamy w następnym rozdziale, przy opisie własności transformacyjnych pola elektromagnetycznego.

Równości, w których obie strony transformują się tak samo pod działaniem każdego z operatorów jakiejś grupy, nazywamy **kowariantnymi** względem tej grupy. Na przykład powyższa definicja czteropędności jest kowariantna względem transformacji Lorentza. Można ją zatem stosować w każdym inercjalnym układzie odniesienia.

Ćwiczenie Pokazać, że równość $a_z = |\mathbf{a}|$ nie jest kowariantna względem grupy obrotów. Można ją stosować, ale tylko w układach, w których oś z jest równoległa do wektora \mathbf{a} .

•••••

Czteropęd cząstki jest zdefiniowany wzorem

$$p^\mu = mu^\mu, \quad (1.52)$$

gdzie m jest masą cząstki⁹, a p^0 jej energią¹⁰. Wobec (1.51), $p^2 = m^2$ i otrzymujemy relatywistyczny wzór na **energię cząstki swobodnej**

⁹Dawniej pisało się masa spoczynkowa, czy skalarna, żeby się odciąć od "masy relatywistycznej", występującej na przykład w słynnym wzorze $E = mc^2$. Teraz przyjęło się, że masa zawsze oznacza masę spoczynkową. Wzór $E = mc^2$ jest odstępstwem od tej konwencji, chyba że ograniczymy się do cząstek w spoczynku.

¹⁰Tę definicję można uwiarygodnić, rozważając zachowanie energii i pędu w elastycznym rozpraszaniu cząstek [9]

$$E^2 = m^2 + \mathbf{p}^2. \quad (1.53)$$

W fizyce klasycznej przyjmuje się jako oczywiste, że samą energię należy liczyć jako pierwiastek ze znakiem plus z wyrażenia po prawej stronie. Zobaczymy dalej, że w mechanice kwantowej sprawa jest bardziej skomplikowana. Zauważmy, że czteropęd jest dobrze określony także dla cząstki o masie zero, poruszającej się z prędkością światła, bo można zrobić przejście graniczne $|\mathbf{v}| \rightarrow 1$ i $m \rightarrow 0$ tak, żeby iloczyn $m\gamma$ dążył do (dowolnej) skończonej granicy. Zakładając, że $\mathbf{p}^2 \ll m^2$ i rozwijając prawą stronę relacji $E = \sqrt{m^2 + \mathbf{p}^2}$ w szereg według potęg małego parametru $\frac{\mathbf{p}^2}{m^2}$, otrzymujemy

$$E = m + \frac{\mathbf{p}^2}{2m} - \frac{(\mathbf{p}^2)^2}{8m^3} + \dots \quad (1.54)$$

Pierwszy wyraz daje identyczność masy cząstki z jej energią spoczynkową, drugi jest znanym z fizyki nierelatywistycznej wyrażeniem na energię kinetyczną cząstki, trzeci daje poprawkę relatywistyczną do energii kinetycznej i będzie potrzebny w paragrafie 5.8 przy obliczaniu relatywistycznych poprawek do poziomów energetycznych atomu wodoru.

Pokażemy jeszcze, że czwórka operatorów $\frac{\partial}{\partial x^\mu}$ transformuje się jak czterowektor z dolnym wskaźnikiem. Z tożsamości

$$dx'^\mu = \frac{\partial x'^\mu}{\partial x^\alpha} dx^\alpha = \Lambda^\mu{}_\alpha dx^\alpha, \quad (1.55)$$

widać, że

$$\frac{\partial x'^\mu}{\partial x^\alpha} = \Lambda^\mu{}_\alpha. \quad (1.56)$$

Definiujemy teraz macierz $\bar{\Lambda}$:

$$\frac{\partial}{\partial x'^\mu} = (\bar{\Lambda})^\alpha{}_\mu \frac{\partial}{\partial x^\alpha}. \quad (1.57)$$

Zauważmy, że

$$\bar{\Lambda}^\alpha{}_\nu = \frac{\partial x^\alpha}{\partial x'^\nu}. \quad (1.58)$$

Pozostaje wykazać, że $\bar{\Lambda} = \Lambda^{-1}$. Rzeczywiście:

$$\Lambda^\mu{}_\alpha \bar{\Lambda}^\alpha{}_\nu = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial x'^\nu} = \frac{\partial x'^\mu}{\partial x'^\nu} = \delta^\mu{}_\nu. \quad (1.59)$$

Często stosuje się oznaczenia

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu}, \quad \partial^\mu \equiv \frac{\partial}{\partial x_\mu}, \quad (1.60)$$

gdzie drugi wzór powstaje z pierwszego przez zmianę poziomu wskaźników po obu stronach.

Prostym zastosowaniem tego wyniku jest wyjaśnienie różnicy w znaku przy kwantowaniu pędu i energii. Tu i w dalszym ciągu używamy **układu jednostek**, w którym zredukowana stała Plancka $\hbar = 1$. W nierelatywistycznej mechanice kwantowej mamy

$$\hat{\mathbf{p}} = -i\nabla. \quad (1.61)$$

Z tego wzoru nie widać, czy przestrzenne składowe pędu powinny być interpretowane jako składowe wektora z wskaźnikami górnymi (p^i), czy z dolnymi ($-p_i$). Odpowiadające tym interpretacjom zerowe składowe są \hat{p}^0 i $-\hat{p}^0$, nie jest więc obojętne, który z tych operatorów jest równy $-i\partial_t$. Od podstawowych wzorów wymagamy, żeby były kowariantne względem transformacji Lorentza. Po prawej stronie mamy składowe przestrzenne wektora ze wskaźnikami dolnymi, więc tak samo należy interpretować składowe \mathbf{p} , i kowariantny wzór ma postać

$$\hat{p}_\mu = +i\frac{\partial}{\partial x^\mu}, \quad (1.62)$$

co oprócz wzoru (1.61) daje poprawnie $\hat{H} = +i\frac{\partial}{\partial t}$. **Równanie Schrödingera zależne od czasu**

$$i\frac{\partial\psi}{\partial t} = \hat{H}\psi, \quad (1.63)$$

jest, przy odpowiedniej definicji hamiltonianu, relatywistycznie kowariantne i ma znacznie szerszy zakres stosowalności niż nierelatywistyczna mechanika kwantowa. Spotkamy je znowu przy omawianiu równania Diraca.

1.9 Przykłady zastosowania czterowektorów i skalarów

Kluczowym punktem w wielu zastosowaniach jest znalezienie skalara, który prosto się wiąże z interesującymi nas wielkościami. Zilustrujemy tę metodę dwoma przykładami.

W układzie laboratorium, cząstka o masie m_1 , pędzie \mathbf{p}_1 i energii $E_1 \equiv E_{Lab}$ zderza się z nieruchomą cząstką o masie m_2 , pędzie $\mathbf{p}_2 = \mathbf{0}$ i energii $E_2 = m_2$. Wyliczmy E_{cms} , czyli całkowitą energię tej pary cząstek w układzie środka masy, w którym z definicji:

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{0}. \quad (1.64)$$

Można by zrobić odpowiednią transformację Lorentza, ale prościej jest zauważyć, że wobec (1.64)

$$s \equiv (p_1 + p_2)^2 = E_{cms}^2. \quad (1.65)$$

Z drugiej strony s , jako skalar, ma tę samą wartość w układzie laboratorium, więc

$$s = p_1^2 + 2p_1p_2 + p_2^2 = m_1^2 + 2m_2E_{Lab} + m_2^2. \quad (1.66)$$

Stąd

$$\sqrt{s} = E_{cms} = \sqrt{2m_2E_{Lab} + m_1^2 + m_2^2} \quad (1.67)$$

Ten wynik tłumaczy popularność zderzaczy wiązek przeciwbieżnych. Na przykład dla zderzeń protonów ($m_1 = m_2 \approx 1 \text{ GeV}$), żeby uzyskać $E_{cms} = 100 \text{ GeV}$,

trzeba by rozpraszać na nieruchomej tarczy protony o energii $E_{Lab} \approx 5000 \text{ GeV}$. Znacznie taniej jest zbudować dwie przeciwbieżne wiązki po 50 GeV każda.

Efekt Döpplera. W punktochwili $x = y = z = t = 0$ znajduje się pojazd jadący wzdłuż osi z z prędkością v (odpowiadającą czynnikiem Lorentza γ) względem układu laboratoryjnego i emitujący w płaszczyźnie (z, x) kwanty światła, które w układzie spoczynkowym pojazdu mają częstość ω_0 . Nieruchomy w układzie laboratoryjnym obserwator odbiera kwanty wyemitowane pod kątem θ ($x/z = \tan \theta$) i mierzy ich częstość ω . Należy znaleźć funkcję $\omega(\omega_0, v, \theta)$.

W układzie spoczynkowym pojazdu, jego czteropędkość $u^\mu = \{1, 0, 0, 0\}$, więc oznaczając czteropęd fotonu przez p_γ mamy

$$u_\mu p_\gamma^\mu = \omega_0. \quad (1.68)$$

W układzie laboratorium ten iloczyn ma tę samą wartość, ale tam

$$u^\mu = \{\gamma, 0, 0, \gamma v\} \quad (1.69)$$

$$p_\gamma^\mu = \{\omega, \omega \sin \theta, 0, \omega \cos \theta\}. \quad (1.70)$$

Porównując wyrażenia na iloczyn otrzymane w obu układach

$$\omega_0 = \gamma\omega - \gamma v\omega \cos \theta \quad (1.71)$$

i rozwiązując na ω mamy:

$$\omega = \frac{\omega_0}{\gamma(1 - v \cos \theta)}. \quad (1.72)$$

Dla $\cos \theta = \pm 1$ otrzymujemy

$$\omega = \omega_0 \frac{\sqrt{1 - v^2}}{1 \mp v} = \omega_0 \sqrt{\frac{1 \pm v}{1 \mp v}}, \quad (1.73)$$

a więc znany wynik, że dla obserwatora stojącego przed pojazdem częstość jest większa, a dla obserwatora stojącego za pojazdem mniejsza (słynne w astronomii przesunięcie ku czerwieni) niż ω_0 . Pod kątem 90° , $\cos \theta = 0$ i otrzymujemy $\omega = \frac{\omega_0}{\gamma}$. To jest poprzeczny efekt Döpplera, przewidziany dopiero przez Einsteina.

1.10 Tensory

Tensory w przestrzeni Minkowskiego definiuje się podobnie jak w przestrzeni euklidesowej, trzeba tylko rozróżniać wskaźniki górne i dolne. Wzór na transformacje Lorentza dla tensora mającego m wskaźników górnych i n dolnych ma postać

$$T_{\nu_1 \dots \nu_n}^{\mu_1 \dots \mu_m} = \left(\prod_{k=1}^m \Lambda^{\mu_k}_{\alpha_k} \right) T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m} \left(\prod_{l=1}^n (\Lambda^{-1})^{\beta_l}_{\nu_l} \right) \quad (1.74)$$

i każdy zbiór liczb, który tak się transformuje jest z definicji tensorem.

Ważne jest, że niektóre stałe macierze są jednocześnie tensorami. Nazywamy je **tensorami niezmienniczymi**. Stała macierz jest tensorem niezmienniczym wtedy i tylko wtedy, kiedy jej iloczyn z odpowiednią liczbą wektorów,

w którym wszystkie wskaźniki są zwięzione, jest skalar. Wykażemy to na przykładzie macierzy $g_{\mu\nu}$, dla której iloczyn z dwoma wektorami $g_{\mu\nu}a^\mu b^\nu$ jest skalar. Uogólnienie na macierze wielowymiarowe, czyli z większą liczbą wskaźników, nie wprowadza żadnych dodatkowych trudności.

Jeżeli $g_{\mu\nu}$ jest tensorem, to iloczyn $g_{\mu\nu}a^\mu b^\nu$, w którym nie ma niezwięzonych wskaźników, musi być skalar. To pokazuje, że warunek jest konieczny. Pozostaje wykazać warunek dostateczny. Ze skalarowości iloczynu wektorów ab i ze wzoru (1.33)

$$a'^\alpha g'_{\alpha\beta} b'^\beta \equiv a^\rho \Lambda^\alpha_\rho g'_{\alpha\beta} \Lambda^\beta_\sigma b^\sigma = a^\rho g_{\rho\sigma} b^\sigma, \quad (1.75)$$

gdzie $g' = g$. Wprowadziliśmy oznaczenie g' , żeby wzory transformacyjne wyglądały jak dla innych tensorów. Ponieważ wszystkie składowe wektorów a i b są dowolne, wynika stąd

$$g_{\rho\sigma} = g'_{\alpha\beta} \Lambda^\alpha_\rho \Lambda^\beta_\sigma. \quad (1.76)$$

Mnożąc to równanie stronami przez $(\Lambda^{-1})^\rho_\mu (\Lambda^{-1})^\sigma_\nu$ i wykonując zwiężenia, otrzymujemy

$$g'_{\mu\nu} = g_{\rho\sigma} (\Lambda^{-1})^\rho_\mu (\Lambda^{-1})^\sigma_\nu, \quad (1.77)$$

a więc poprawny wzór na transformację tensora z dwoma dolnymi wskaźnikami. Jak widać, jest wszystko jedno, czy założymy, że $g_{\mu\nu}$ nie zmienia się przy transformacji Lorentza, czy też że transformuje się według wzoru (1.74) jak tensor z dwoma dolnymi wskaźnikami. Tensor metryczny $g_{\mu\nu}$ jest więc zgodnie z przyjętą definicją tensorem, ale jednocześnie jest stałą macierzą. To znaczy, że jest tensorem niezmienniczym. Z faktu, że $g_{\mu\nu}$ **jest tensorem niezmienniczym** wynika, że stosując zwykłe reguły podnoszenia wskaźników otrzymujemy kolejne tensory niezmiennicze, z czego już korzystaliśmy w paragrafie 1.7.

Ćwiczenie Korzystając ze wzoru (1.33) pokazać, że macierz Λ jest tensorem niezmienniczym. Przedyskutować z tego punktu widzenia wzory (1.17).

•••••

Pokażemy teraz, że **macierz całkowicie antysymetryczna** $\epsilon_{\alpha\beta\gamma\delta}$ jest tensorem niezmienniczym względem właściwych transformacji Lorentza. Definicja tego tensora jest analogiczna do definicji (pseudo)tensora całkowicie antysymetrycznego w geometrii euklidesowej. Element macierzowy $\epsilon_{\alpha\beta\gamma\delta} = 0$, jeśli którekolwiek dwa z jego czterech wskaźników są sobie równe. Dla niezerowych elementów, ciąg wskaźników $\alpha, \beta, \gamma, \delta$ musi więc być permutacją liczb 0, 1, 2, 3. Jeżeli ta permutacja jest parzysta, to $\epsilon_{\alpha\beta\gamma\delta} = 1$, jeśli jest nieparzysta, $\epsilon_{\alpha\beta\gamma\delta} = -1$. Na mocy tej definicji,

epsdet

$$D \equiv \epsilon_{\alpha\beta\gamma\delta} a^\alpha b^\beta c^\gamma d^\delta = \begin{vmatrix} a^0 & a^1 & a^2 & a^3 \\ b^0 & b^1 & b^2 & b^3 \\ c^0 & c^1 & c^2 & c^3 \\ d^0 & d^1 & d^2 & d^3 \end{vmatrix}, \quad (1.78)$$

gdzie wyrażenie po prawej stronie jest wyznacznikiem zbudowanym ze składowych czterowektorów a, b, c, d . Żeby wykazać, że $\epsilon_{\alpha\beta\gamma\delta}$ jest tensorem niezmienniczym, trzeba udowodnić, że wyznacznik D jest niezmiennikiem.

Ponieważ transpozycja nie zmienia wartości wyznacznika, a pomnożenie wszystkich wyrazów w jednym wierszu przez -1 zmienia jego znak, zachodzi równość

$$D = - \begin{vmatrix} a^0 & b^0 & c^0 & d^0 \\ -a^1 & -b^1 & -c^1 & -d^1 \\ -a^2 & -b^2 & -c^2 & -d^2 \\ -a^3 & -b^3 & -c^3 & -d^3 \end{vmatrix}. \quad (1.79)$$

Mnożąc macierz tego wyznacznika lewostronnie przez macierz wyznacznika ze wzoru (1.78) i korzystając z twierdzenia, że wyznacznik z iloczynu macierzy jest iloczynem wyznaczników tych macierzy mamy

$$-D^2 = \begin{vmatrix} a^2 & ab & ac & ad \\ ba & b^2 & bc & bd \\ ca & cb & c^2 & cd \\ da & db & dc & d^2 \end{vmatrix}, \quad (1.80)$$

a więc skalar. Z faktu, że D^2 jest skalar, że przy transformacji Lorentza D albo się nie zmienia, albo zmienia znak. To już dowodzi, że dla $D^2 = 0$, wyznacznik D jest niezmiennikiem. Pozostaje do rozważenia przypadek $D^2 \neq 0$. Ograniczmy się teraz do właściwych transformacji Lorentza. Przy takich transformacjach składowe wektorów a, b, c, d , a więc i wyznacznik D , zmieniają się w sposób ciągły. Nie może więc być przeskoku między wartościami $\pm\sqrt{-D^2}$. Wnioskujemy stąd, że D nie zmienia się przy właściwych transformacjach Lorentza, a zatem że $\epsilon_{\alpha\beta\gamma\delta}$ **jest tensorem niezmienniczym** względem takich transformacji.

Ćwiczenie Pokazać, że element objętości w czasoprzestrzeni d^4x jest skalar, i że to samo zachodzi dla elementu objętości w przestrzeni pędów i energii d^4p .

•••••

Przy odbiciu przestrzennym, przestrzenne składowe czterowektorów zmieniają znaki, a składowe czasowe pozostają bez zmian. W tych warunkach wyznacznik D zmienia znak. Podobnie dzieje się przy odbiciu czasowym, przy którym zmieniają znaki czasowe składowe wektorów, a składowe przestrzenne pozostają bez zmiany. Wyznacznik D , a co zatem idzie i $\epsilon_{\alpha\beta\gamma\delta}$, nie zmienia się przy właściwych transformacjach Lorentza i przy transformacjach Lorentza z odbiciem zupełnym. Względem takich transformacji ϵ jest więc tensorem niezmienniczym. Jest natomiast pseudotensorem, jeśli włączyć odbicia przestrzenne i czasowe.

Ćwiczenie Stosując metodę z poprzedniego dowodu pokazać, że w euklidesowej przestrzeni trójwymiarowej tensor całkowicie antysymetryczny $\epsilon_{\alpha\beta\gamma}$ jest niezmienniczy względem obrotów.

•••••

W zastosowaniach często trzeba całkować po pędzie cząstki. Załóżmy, że w układzie spoczynkowym cząstki $d^3p = dp_x dp_y dp_z$. Jeśli przejdziemy do innego układu odniesienia, transformując odpowiednio dp_x, dp_y, dp_z , to na ogół otrzymamy inne d^3p , co czasem utrudnia rachunki. Rozważmy najpierw pchnięcia

bezobrotowe wzdłuż osi z . Dla pędu cząstki można położyć $p_z = m_T \sinh y$, $E(\mathbf{p}) = m_T \cosh y$, gdzie $m_T = \sqrt{m^2 + p_x^2 + p_y^2}$. W takim razie

$$d^3p = dp_x dp_y m_T \cosh y dy = E(\mathbf{p}) dp_x dp_y dy \quad (1.81)$$

Parametr y jest równy pośpieszności układu, w którym cząstka spoczywa, względem układu, w którym określony jest pęd \mathbf{p} . Przy pchnięciu wzdłuż osi z , dp_x i dp_y nie zmieniają się. Zgodnie z uwagą na końcu paragrafu 1.4, nie zmienia się też dy . Wynika stąd, że wyrażenie

$$d\Omega \equiv \frac{d^3p}{E(\mathbf{p})} \quad (1.82)$$

nie zmienia się przy pchnięciach wzdłuż osi z . Ponieważ oś z jest wybrana dowolnie, wynik uogólnia się na wszystkie pchnięcia bezobrotowe. Widać też, że element (1.82) nie zmienia się przy obrotach. Wynika stąd, że jest niezmiennikiem względem wszystkich właściwych transformacji Lorentza. Przypuśćmy, że całka

$$\int d^3p f(\mathbf{p}) = \int \frac{d^3p}{E(\mathbf{p})} E(\mathbf{p}) f(\mathbf{p}) \quad (1.83)$$

jest skalarem. W wyrażeniu po lewej stronie ani d^3p , ani $f(\mathbf{p})$ nie są skalarami. Ważne jest więc, żeby oba liczyć w tym samym układzie odniesienia. W wyrażeniu po prawej stronie zarówno $E(\mathbf{p})f(\mathbf{p})$ jak i $\frac{d^3p}{E(\mathbf{p})}$ są skalarami, a zatem nie zależą od tego, w jaki układzie je liczymy. To często ułatwia rachunki.

Ćwiczenie Pokazać, że pierwszy z czynników $E(\mathbf{p})$ we wzorze (1.83) poprawia na skrócenie Lorentza, a drugi na związany z nim wzrost gęstości.

•••••

Ćwiczenie Pokazać, że wyrażenie $\delta(p^2 - m^2)d^4p$ jest skalarem i że scałkowane po p^0 daje, z dokładnością do czynnika 2, wyrażenie (1.82). Dlaczego całkowanie po p^0 nie psuje niezmienniczości?

•••••

1.11 Uzupełnienie: Liniowość transformacji Lorentza

Dla uzupełnienia uzasadnienia wzoru (1.17) pokażemy teraz, że transformacje Lorentza muszą być liniowe. Różniczkując równość (1.8) stronami po x^γ , otrzymujemy

$$\frac{\partial^2 x'^\mu}{\partial x^\alpha \partial x^\gamma} g_{\mu\nu} \frac{\partial x'^\nu}{\partial x^\beta} + \frac{\partial x'^\mu}{\partial x^\alpha} g_{\mu\nu} \frac{\partial^2 x'^\nu}{\partial x^\beta \partial x^\gamma} = 0. \quad (1.84)$$

Ta równość jest prawdziwa dla dowolnych α, β, γ . Można więc zmienić α na γ , γ na β i β na α , co daje

$$\frac{\partial^2 x'^\mu}{\partial x^\gamma \partial x^\beta} g_{\mu\nu} \frac{\partial x'^\nu}{\partial x^\alpha} + \frac{\partial x'^\mu}{\partial x^\gamma} g_{\mu\nu} \frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\beta} = 0. \quad (1.85)$$

Odejmując to równanie od poprzedniego i wykorzystując symetrię macierzy g dostajemy

$$\frac{\partial^2 x'^\mu}{\partial x^\alpha \partial x^\gamma} g_{\mu\nu} \frac{\partial x'^\nu}{\partial x^\beta} - \frac{\partial x'^\mu}{\partial x^\gamma} g_{\mu\nu} \frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\beta} = 0. \quad (1.86)$$

Wymieniając w tej równości wskaźniki α i γ otrzymujemy równość

$$\frac{\partial x'^\mu}{\partial x^\alpha} g_{\mu\nu} \frac{\partial^2 x'^\nu}{\partial x^\beta \partial x^\gamma} - \frac{\partial^2 x'^\mu}{\partial x^\alpha \partial x^\gamma} g_{\mu\nu} \frac{\partial x'^\nu}{\partial x^\beta} = 0, \quad (1.87)$$

która odjęta stronami od równości (1.84) daje, po podzieleniu obu stron równania przez dwa,

$$\frac{\partial^2 x'^\mu}{\partial x^\alpha \partial x^\gamma} g_{\mu\nu} \frac{\partial x'^\nu}{\partial x^\beta} = 0. \quad (1.88)$$

Przy każdym wyborze wskaźników α i γ jest to liniowy jednorodny układ czterech równań ($\beta = 0, 1, 2, 3$) na cztery niewiadome $\frac{\partial^2 x'^\mu}{\partial x^\alpha \partial x^\gamma}$, $\mu = 0, 1, 2, 3$, którego wyznacznik główny jest różny od zera (prosty wniosek z $\text{Det}g \neq 0$ i (1.9)), a zatem jedynym rozwiązaniem jest

$$\frac{\partial^2 x'^\mu}{\partial x^\alpha \partial x^\gamma} = 0. \quad (1.89)$$

Znikanie wszystkich drugich pochodnych oznacza, że transformacja jest liniowa, co było do udowodnienia.

Rozdział 2

Cząstka w polu elektromagnetycznym

2.1 Przypomnienie wiadomości z klasycznej elektrodynamiki

Podstawowymi pojęciami w klasycznej elektrodynamice są wektory pola. Pole elektryczne \mathbf{E} jest wektorem, a pole magnetyczne \mathbf{B} pseudowektorem, w trójwymiarowej przestrzeni euklidesowej. Jako pojęcie pomocnicze wprowadza się niekiedy **potencjał wektorowy** A^μ , który jest czterowektorem. Wektory pola są związane z potencjałem wektorowym wzorami

vecpol

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad \mathbf{E} = -\nabla A^0 - \frac{\partial}{\partial t} \mathbf{A}, \quad (2.1)$$

ale bez potencjału wektorowego można się obejść, bo zarówno równania Maxwella, jak i wzór Lorentza na siłę wywieraną przez pole na naładowaną cząstkę punktową, dadzą się zapisać z pomocą samych tylko pól. Zobaczymy, że w mechanice kwantowej jest inaczej. Zauważmy, że dla niezależnego od czasu potencjału wektorowego, dla którego $\mathbf{A}(x) \equiv \mathbf{0}$, A^0 jest potencjałem znanym z elektrostatyki.

Z potencjału wektorowego można zbudować tensor antysymetryczny, znany jako **tensor pola**:

tenpol

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (2.2)$$

Ważne jest, że jak łatwo sprawdzić przez podstawienie, tensor pola nie zmienia się, jeśli poddamy potencjał wektorowy **transformacji cechowania**

trcech

$$A^\mu(x) \rightarrow A'^\mu(x) = A^\mu(x) + \partial^\mu \Lambda(x), \quad (2.3)$$

gdzie $\Lambda(x)$ jest jakąkolwiek (mającą wszystkie drugie pochodne ciągłe) skalarną funkcją składowych czterowektora x . Zauważmy, że dla $\mu = 1, 2, 3$ operator ∂^μ jest równy $-\frac{\partial}{\partial x^\mu}$, a więc

trcec3

$$\mathbf{A}'(x) = \mathbf{A}(x) - \nabla \Lambda(x). \quad (2.4)$$

Korzystając ze wzorów (2.1) otrzymujemy

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} = -F^{\nu\mu}. \quad (2.5)$$

Ćwiczenie Pokazać, że pod wpływem bezobrotowego pchnięcia z prędkością v wzdłuż osi z , pole elektryczne i pole magnetyczne transformują się według wzorów:

$$\begin{aligned} E'_x &= \gamma(E_x + vB_y), & E'_y &= \gamma(E_y - vB_x), & E'_z &= E_z, \\ B'_x &= \gamma(B_x - vE_y), & B'_y &= \gamma(B_y + vE_x), & B'_z &= B_z. \end{aligned}$$

•••••

Wektory \mathbf{E} i \mathbf{B} są zwykłymi wektorami euklidesowymi, więc poziom wskaźników nie ma dla nich znaczenia. Niezmienniczość tensora pola wobec cechowania oznacza, że pole elektromagnetyczne (\mathbf{E}, \mathbf{B}) nie zależy od wycechowania (to znaczy od wyboru funkcji $\Lambda(x)$). Danemu polu elektromagnetycznemu odpowiada nieskończenie wiele różnych potencjałów wektorowych, ale całka

$$\oint_C ds \mathbf{A} = \iint_{\Sigma} d\Sigma \nabla \times \mathbf{A} = \Phi_{\Sigma}(\mathbf{B}) \quad (2.6)$$

jest niezmiennicza względem zmian wycechowania. W tym wzorze C jest krzywą zamkniętą, w trójwymiarowej przestrzeni euklidesowej, stanowiącą brzeg powierzchni Σ^1 . Pierwsza równość wyraża więc znane z analizy matematycznej twierdzenie Stokesa. Założenia twierdzenia Stokesa przypomnimy na końcu następnego paragrafu. Funkcja $\Phi_{\Sigma}(\mathbf{B})$ oznacza strumień pola magnetycznego przez powierzchnię Σ , więc druga równość wynika z pierwszego ze wzorów (2.1).

Pierwsze cztery równania Maxwella można zapisać w postaci

$$\partial_{\mu} F^{\mu\nu} = 4\pi J^{\nu}, \quad (2.7)$$

gdzie zerowa składowa **czterowektora prądu** $J = \{\rho, \mathbf{j}\}$ jest gęstością ładunku, a pozostałe są składowymi gęstości prądu. Działając na lewą stronę tego równania operatorem ∂_{ν} otrzymujemy

$$\partial_{\nu} \partial_{\mu} F^{\mu\nu} = 0. \quad (2.8)$$

Wyrażenie po lewej stronie, jako wynik zwięzienia symetrycznego operatora $\partial^{\nu} \partial^{\mu}$ z antysymetrycznym tensorem $F^{\mu\nu}$, jest identycznie równe zero. Wynika stąd **równanie ciągłości**

$$\partial_{\nu} J^{\nu} \equiv \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (2.9)$$

To równanie oznacza, że przyrost ładunku w danej objętości jest równy ilości ładunku wniesionego do tej objętości przez prąd. Ładunek jest więc zachowywany ani nie ginie, ani nie powstaje. Pozostałe cztery równania Maxwella są

¹Żeby całka po C była określona, musi być wybrany kierunek obiegu na tej krzywej. Krzywa C jest brzegiem powierzchni Σ , jeśli obiega ją w kierunku dodatnim (przeciwym do kierunku wskazówek zegara).

tożsamościowo spełnione, jeśli wyrażamy pola przez potencjał wektorowy według wzorów (2.1).

Ćwiczenie Sprawdzić, że równania Maxwella

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0.$$

wynikają z równości (2.1).

•••••

Wzór na siłę Lorentza można zapisać w następującej, jawnie kowariantnej postaci

$$\frac{dp^\mu}{d\tau} = qF^{\mu\nu}u_\nu, \quad (2.10)$$

gdzie τ jest czasem własnym cząstki, q jej ładunkiem i u_ν czteropędkością.

Ćwiczenie Pokazać, że w układzie spoczynkowym cząstki $\frac{d\mathbf{p}}{dt} = q\mathbf{E}$. Czy zakładając tylko ten wzór, można otrzymać ogólny wzór na siłę Lorentza?

Wskazówka: rozpatrzeć przypadek, kiedy w układzie spoczynkowym $\mathbf{E} = \mathbf{0}, \mathbf{B} \neq \mathbf{0}$.

•••••

Stwierdzonym doświadczalnie elektromagnetycznym zjawiskiem, które nie da się wytłumaczyć przy pomocy klasycznej elektrodynamiki, jest efekt Bohma-Aharonova.

2.2 Efekt Bohma-Aharonova

Rozważamy cząstki, z których każda wychodzi ze wspólnego źródła i może dolecieć do ekranu po dwu drogach C_+ i C_- . W tych warunkach, podobnie jak w optyce, na ekranie powstają prążki interferencyjne. Między drogi C_+ i C_- wstawiamy walec, którego ścianki są nieprzenikalne dla pola magnetycznego. W praktyce można to osiągnąć, na przykład, pokrywając je warstwą nadprzewodnika. Drogi cząstek znajdują się po obu stronach walca, całkowicie na zewnątrz. Na drogach C_\pm i w ich otoczeniu pole magnetyczne jest równe zero. Natomiast wewnątrz walca jest jednorodne pole magnetyczne \mathbf{B} równoległe do osi. Można sobie wyobrazić, że tory C_+ i C_- leżą w jednej płaszczyźnie, prostopadłej do osi walca. Ponieważ cząstki poruszają się w obszarze, gdzie natężenie pola magnetycznego jest równe zero, według klasycznej fizyki ich ruch nie powinien zależeć od natężenia pola \mathbf{B} wewnątrz walca. To przewidywanie jest sprzeczne z doświadczeniem. Przy zmianach natężenia pola \mathbf{B} obserwuje się przesunięcia prążków – zjawisko znane jako **efekt Bohma-Aharonova**.

Efekt Bohma-Aharonova można by zrozumieć, gdyby przyjąć, sprzecznie z klasyczną elektrodynamiką, że potencjał wektorowy wpływa na ruch cząstki nie tylko poprzez pole elektromagnetyczne. Trzeba jednak najpierw sprawdzić, czy nie da się dobrać wycechowania tak, żeby w otoczeniu torów cząstek zachodziło $\mathbf{A} = \mathbf{0}$, niezależnie od natężenia pola wewnątrz walca, bo wtedy i oddziaływanie z potencjałem wektorowym nic by nie pomogło. Interesują nas tu tylko składowe przestrzenne potencjału wektorowego, bo tylko one wiążą się z natężeniem pola magnetycznego. Sprawdzimy więc, kiedy potencjał wektorowy, w jakimś

obszarze obejmującym tory cząstek C_{\pm} , da się przeprowadzić przez zmianę wycechowania w **potencjał, dla którego $\mathbf{A} = \mathbf{0}$** , niezależnie od natężenia pola magnetycznego wewnątrz walca. W takim przypadku, zmiana pola magnetycznego wewnątrz walca nie prowadziłyby do zmiany, wzdłuż torów cząstek, nie tylko pola \mathbf{B} , ale i związanego z tym polem potencjału wektorowego.

Znajdziemy ogólne warunki, przy których potencjał wektorowy A'^{μ} można przeprowadzić, robiąc transformację cechowania, w potencjał A^{μ} , dla którego $\mathbf{A} = \mathbf{0}$. Ponieważ, niezależnie od wycechowania, $\nabla \times \mathbf{A} = \mathbf{B}$, warunkiem koniecznym jest $\mathbf{B} = \mathbf{0}$. Warunkiem wystarczającym jest istnienie takiej skalarnej funkcji $\Lambda(x)$, że

$$\mathbf{A}'(x) = -\nabla\Lambda(x). \quad (2.11)$$

Wtedy na mocy wzoru (2.4) mamy $\mathbf{A}(x) \equiv \mathbf{0}$. Zauważmy, że powyższy warunek jest taki sam, jak warunek, który w mechanice musi spełniać pole sił, odpowiednik pola wektorowego $\mathbf{A}'(x)$, żeby istniał potencjał, odpowiednik skalarne go pola $\Lambda(x)$. Potencjał w mechanice istnieje, jeśli praca sił wzdłuż każdej drogi zamkniętej jest równa zero. Analogicznie, wycechowanie w którym $\mathbf{A} \equiv \mathbf{0}$ istnieje, jeśli dla każdej krzywej zamkniętej w badanym obszarze całka okrężna $\oint ds \mathbf{A}' = 0$. Wybierzmy krzywą zamkniętą C i oznaczmy powierzchnię stanowiącą jej wnętrze Σ . Jeśli $\mathbf{B} = \mathbf{0}$ na Σ i C , to znikanie całki okrężnej wynika z twierdzenia Stokesa, ale pod dwoma warunkami. Po pierwsze, funkcja $\mathbf{A}'(x)$ musi być dostatecznie regularna², co w szczególności znaczy, że nie ma osobliwości na Σ . Po drugie, każda krzywa zamknięta na Σ powinna dać się ściągnąć do punktu nie wychodząc poza Σ . To znaczy, że nie może być dziur w powierzchni Σ . Jeśli funkcja $\Lambda(x)$ spełniająca warunek (2.11) nie istnieje, to nie istnieje też wycechowanie, w którym $\mathbf{A}(x) \equiv \mathbf{0}$.

Żeby zastosować te wyniki do efektu Bohma-Aharonova, rozważmy krzywą zamkniętą C złożoną z krzywej C_- i krzywej C_+ . Kierunek obiegu krzywej C jest na jednej z krzywych C_{\pm} zgodny z kierunkiem prędkości cząstki, a na drugiej przeciwny. Wybieramy go tak, żeby krzywa C była brzegiem jakiejś powierzchni Σ przecinającej się z walcem. Całka okrężna z potencjału wektorowego po krzywej C jest równa strumieniowi pola magnetycznego przez dowolną powierzchnię dla której C jest brzegiem, więc przy $\mathbf{B} \neq \mathbf{0}$ otrzymujemy niezerowy strumień i co za tym idzie niezerową wartość całki okrężnej. Tak więc potencjał wektorowy nie da się przez zmianę wycechowania sprowadzić do postaci, dla której $\mathbf{A}(x) \equiv \mathbf{0}$ na obu drogach C_{\pm} . Podane tu rozumowanie prowadzi też do mocniejszego wniosku, że potencjał $\mathbf{A}(x)$ zależy od natężenia pola \mathbf{B} wewnątrz walca.

Ćwiczenie Usuając z powierzchni Σ część znajdującą się wewnątrz walca, otrzymujemy powierzchnię Σ' , na której $\mathbf{B} \equiv \mathbf{0}$. Wyjaśnić dlaczego, mimo to, nie można wnioskować, że $\mathbf{A}(x) \equiv \mathbf{0}$ na całej krzywej C .

•••••

Zauważmy jeszcze, że da się wybrać obszar, bez dziur, zawierający krzywą C_+ , w którym $\mathbf{B} = \mathbf{0}$. Można więc wybrać wycechowanie, w którym $\mathbf{A}(x) = \mathbf{0}$ na całej drodze C_+ i w pewnym jej otoczeniu. W tym wycechowaniu zmiana

²Wymagana jest ciągłość składowych wektora \mathbf{A}' na powierzchni Σ i ciągłość ich pierwszych pochodnych cząstkowych na krzywej C .

fazy funkcji falowej cząstki poruszającej się po drodze C_+ nie zależy od pola magnetycznego \mathbf{B} . Zależna od \mathbf{B} różnica faz odpowiedzialna za przesuwanie się prążków pochodzi wtedy wyłącznie od wpływu potencjału wektorowego na drodze C_- . Podobnie, można wybrać inne wycechowanie, w którym $\mathbf{A}(x) = \mathbf{0}$ na drodze C_- i w jej otoczeniu, a więc cały efekt pochodzi od cząstek poruszających się po drodze C_+ . To znaczy, że zmiana fazy na danej drodze jest niefizyczna. Sens fizyczny ma tylko różnica faz nabytych przez cząstkę poruszającą się po dwu różnych drogach, dająca się wyznaczyć doświadczalnie z położenia prążków interferencyjnych.

Ćwiczenie Pokazać, że można wybrać wycechowanie tak, że cały efekt pochodzi od dowolnie małego odcinka którejkolwiek z dróg C_{\pm} .

•••••

2.3 Symetria względem cechowania w mechanice kwantowej

W mechanice kwantowej mamy symetrię względem cechowania, jeśli dla dowolnej rzeczywistej, skalarnej funkcji $\Lambda(x)$ stan cząstki o ładunku q , opisywany przez funkcję falową $\psi(x)$, jest identyczny z jej stanem opisywanym przez funkcję falową

gaugef

$$\psi'(x) = e^{-iq\Lambda(x)}\psi(x). \quad (2.12)$$

Zakładamy tu oczywiście, że funkcja $\Lambda(x)$ jest na tyle regularna, że $\psi'(x)$ jest dopuszczalną funkcją falową. Nazywamy taką symetrię lokalną, bo współczynnik $e^{-iq\Lambda(x)}$ jest na ogół różny w różnych punktachwilach x . Symetria globalna, kiedy $\Lambda(x) = \text{const}$, jest ogólną, dobrze znaną własnością funkcji falowych i nie będziemy jej tu osobno dyskutować. Łatwo sprawdzić, że transformacje (2.12) stanowią grupę. Ponieważ liczby $e^{-iq\Lambda(x)}$ mogą być interpretowane jako macierze unitarne o wymiarze 1×1 , oznaczamy tę grupę $U(1)$. Pokażemy, że założenie lokalnej symetrii cechowania $U(1)$ sugeruje pewne oddziaływanie cząstek naładowanych z polem elektromagnetycznym. Innym pytaniem jest, czy akurat to oddziaływanie prowadzi do wyników zgodnych z doświadczeniem. Taki sposób wprowadzania oddziaływań odgrywa ważną rolę we współczesnej fizyce. Dla wielokomponentowych funkcji falowych można za $\Lambda(x)$ wstawiać macierze. Jeżeli wybierzemy dwuwymiarowe macierze unitarne o wyznaczniku równym jeden (grupa $SU(2)$), otrzymamy znaczną część teorii oddziaływań elektroslabych. Jeżeli wybierzemy trójwymiarowe macierze unitarne o wyznaczniku równym jeden (grupa $SU(3)$), otrzymamy chromodynamikę, czyli teorię oddziaływań silnych.

Przyjęcie symetrii (2.12) wymaga zmian w teorii. Na przykład, kładąc $\Lambda(x) = -\frac{px}{q}$ możemy zmienić funkcję falową $\psi(x) = e^{-ipx}$ w funkcję $\psi'(x) = 1$. Jeżeli ta zmiana ma nie wpływać na czteropęd cząstki, to trzeba zmienić operator czteropędu. Kowariantne względem zmian wycechowania operatory odpowiadające składowym czteropędu muszą spełniać warunki

covmom

$$(\hat{p}^\mu \psi(x))' = \hat{p}'^\mu(x) \psi'(x), \quad (2.13)$$

bo przechodząc do układu primowanego powinniśmy mieć swobodę wyboru, czy transformujemy iloczyn $\hat{p}^\mu \psi(x)$, czy każdy jego czynnik z osobna. Zakładamy, że

$$\hat{p}^\mu \psi(x) = iD^\mu(x)\psi(x). \quad (2.14)$$

pochko

Zwykła pochodna została tu zastąpiona przez **pochodną kowariantną** zdefiniowaną wzorem³

$$D^\mu(x) \equiv \partial^\mu + iqA^\mu(x), \quad (2.15)$$

gdzie potencjał wektorowy A^μ jest czterowektorem, który przy zmianach wycechowania transformuje się tak, żeby czteropęd cząstki spełniał warunki (2.13). Uzasadnimy teraz współczynniki i i q . Współczynnik i jest potrzebny, żeby nowo zdefiniowany operator pędu iD_μ był hermitowski. Współczynnik rzeczywisty – tu wybrany jako ładunek cząstki q – z punktu widzenia symetrii mógłby być jakąkolwiek stałą różną do zera, ale jak wyniknie we wzoru (2.34), przyjęty tu wybór jest jedynym, przy którym interpretując A^μ jako potencjał wektorowy znany z klasycznej elektrodynamiki, odtwarza się poprawnie równanie Schrödingera dla cząstki w polu elektrostatycznym.

covder

Warunki (2.13) dają

$$(D^\mu \psi(x))' = D'^\mu \psi'(x), \quad (2.16)$$

czyli

$$e^{-iq\Lambda(x)}(\partial^\mu + iqA^\mu(x))\psi(x) = (\partial^\mu + iqA'^\mu(x))e^{-iq\Lambda(x)}\psi(x). \quad (2.17)$$

gaugea

Po redukcjach i uporządkowaniu otrzymujemy stąd jako wniosek:

$$A'^\mu(x) = A^\mu(x) + \partial^\mu \Lambda(x). \quad (2.18)$$

W poprzednim paragrafie, taki sam wzór dla potencjału wektorowego wynikał z założenia, że stan fizyczny układu zależy tylko od pól. Tu nie robimy tego założenia, o którym wiemy z dyskusji efektu Bohma-Aharonova, że jest błędne. Według nowej interpretacji dobrze określony jest pęd niesiony przez cząstkę i pole. W chwili $t = 0$, w wycechowaniu $\{\psi(x) = e^{i\mathbf{P}\mathbf{x}}, A = 0\}$ cząstka niesie pęd \mathbf{p} , a pole pęd zero. W wycechowaniu $\{\psi(x) = 1, \mathbf{A}' = \frac{1}{q}\mathbf{p}, A'^0 = 0\}$ cząstka niesie pęd zero, a pole A pęd \mathbf{p} . Te dwie sytuacje są fizycznie nierozróżnialne i odpowiadają temu samemu stanowi układu. Mamy więc symetrię. Biorąc sprzężenie zespolone wzoru (2.16) łatwo sprawdzić, że kowariantnym odpowiednikiem wyrażenia $\partial_\mu \psi^*(x)$ jest $D_\mu^* \psi^*(x)$.

Jako przykład zastosowania przedyskutujemy efekt Bohma-Aharonova. Rozważmy ruch cząstki w płaszczyźnie $\{x, y\}$, używając współrzędnych biegunowych. W kole $r \leq \frac{1}{2}$ jest jednorodne pole magnetyczne \mathbf{B} , równoległe do osi z . Dla $r > \frac{1}{2}$, $\mathbf{B} = 0$. Strumień pola magnetycznego przez płaszczyznę $\{x, y\}$ wynosi więc

$$\Phi(B) = \frac{1}{4}\pi B. \quad (2.19)$$

potwek

Jednym z nieskończenie wielu potencjałów wektorowych opisujących to pole elektromagnetyczne w płaszczyźnie $\{x, y\}$ jest

³Pochodna kowariantna ma prostą interpretację geometryczną, której bardzo dobry opis można znaleźć w monografii [6].

$$A = \left\{0, -\frac{Br}{2} \sin \phi, \frac{Br}{2} \cos \phi, 0\right\}, \quad \text{dla} \quad r \leq \frac{1}{2}, \quad (2.20)$$

$$A = \left\{0, -\frac{B}{8r} \sin \phi, \frac{B}{8r} \cos \phi, 0\right\}, \quad \text{dla} \quad r \geq \frac{1}{2}. \quad (2.21)$$

Cząstka wylatuje ze źródła w punkcie $\{r = 1, \phi = \pi\}$ i dociera do ekranu w punkcie $\{r = 1, \phi = 0\}$. Możliwe są dwie drogi: C_+ z $0 < \phi < \pi$ i C_- z $\pi < \phi < 2\pi$. Dla obu dróg przyjmujemy $r \equiv 1$. Różniczkując funkcje $x = r \cos \phi$ i $y = r \sin \phi$ otrzymujemy element drogi:

$$ds = \{0, -\sin \phi d\phi, \cos \phi d\phi, 0\}. \quad (2.22)$$

Ten element jest na krzywej C_+ , gdzie $d\phi < 0$, antyrównoległy, a na krzywej C_- , gdzie $d\phi > 0$, równoległy, do wektora \mathbf{A} . Długość każdej z dróg wynosi π , mamy więc

$$\int_{C_{\pm}} \mathbf{A} ds = \mp \frac{\pi B}{8} = \mp \frac{1}{2} \Phi(B). \quad (2.23)$$

Potrzebne nam będą jeszcze dwa inne wycechowania, takie że części przestrzenne odpowiadających im potencjałów wektorowych A_{\pm} znikają odpowiednio na drogach C_{\pm} . Na mocy tej definicji i poprzedniego wzoru

bomaha

$$\int_{C_{\pm}} \mathbf{A}_{\pm}(x) ds = 0, \quad \int_{C_{\pm}} (\mathbf{A}_{\pm}(x) - \mathbf{A}(x)) ds = \pm \frac{1}{2} \Phi(\mathbf{B}). \quad (2.24)$$

Istnienie tych wycechowań wynika z dyskusji w poprzednim paragrafie. Oznaczamy Δ_{\pm}^0 zmiany fazy funkcji falowej cząstki przy przejściu od źródła do ekranu, odpowiednio dla dróg C_{\pm} i wycechowań odpowiadających potencjałom wektorowym A_{\pm} . Wielkości Δ_{\pm}^0 nie zależą od pola magnetycznego, bo na drogach C_{\pm} zarówno pole \mathbf{B} jak i odpowiednie potencjały wektorowe A_{\pm} są identycznie równe zero, a więc nie ma tam informacji o polu \mathbf{B} wewnątrz walca.

Ponieważ potencjały \mathbf{A}_{\pm} różnią się od potencjału \mathbf{A} tylko wyborem wycechowania, muszą istnieć takie skalarne funkcje Λ_{\pm} , że

$$\mathbf{A} = \mathbf{A}_{\pm} - \nabla \Lambda_{\pm}. \quad (2.25)$$

Mamy więc

$$\int_{C_{\pm}} (\mathbf{A}_{\pm}(x) - \mathbf{A}(x)) ds = \int_{C_{\pm}} \nabla \Lambda_{\pm} ds = \Delta \Lambda_{\pm} \quad (2.26)$$

gdzie $\Delta \Lambda_{\pm}$ oznaczają przyrosty funkcji Λ_{\pm} przy przejściu od źródła do ekranu po drodze odpowiednio C_+ lub C_- . Na mocy wzoru (2.24)

$$\Delta \Lambda_{\pm} = \pm \frac{1}{2} \Phi(\mathbf{B}) \quad (2.27)$$

Przyrostom $\Delta \Lambda_{\pm}$ odpowiadają przyrosty fazy funkcji falowej odpowiednio o $-q \Delta \Lambda_{\pm}$.

Teraz możemy już obliczyć, w wycechowaniu odpowiadającym potencjałowi wektorowemu \mathbf{A} , pełne przyrosty faz dla funkcji falowych cząstek poruszających się po drogach C_{\pm} . Oznaczmy te przyrosty Δ_{\pm} . Otrzymujemy

$$\Delta_{\pm} = \Delta_{\pm}^0 \mp \frac{1}{2}q\Phi(\mathbf{B}), \quad (2.28)$$

gdzie pierwszy człon daje niezależne od pola magnetycznego przyrosty fazy w wycechowaniach odpowiadających potencjałom A_{\pm} , a drugi poprawki na przejście do wycechowania odpowiadającego potencjałowi A . Różnica faz, która decyduje o położeniu prążków na ekranie, wynosi więc

$$\Delta \equiv \Delta_+ - \Delta_- = \Delta_+^0 - \Delta_-^0 - q\Phi(B). \quad (2.29)$$

Ta różnica faz zmienia się o całkowitą wielokrotność liczby 2π , co nie wpływa na funkcję falową, kiedy strumień pola magnetycznego zmienia się o

$$\Delta\Phi(B) = \frac{2\pi n}{q}, \quad (2.30)$$

gdzie n jest dowolną liczbą całkowitą.

Ćwiczenie Pokazać, że powyższe wyniki nie ulegają zmianie, jeśli drogi C_{\pm} łączące źródło z ekranem przejdą w sposób ciągły w inne drogi C'_{\pm} nie wychodząc poza obszar $\mathbf{B} = 0$.

•••••

Ćwiczenie Wyjaśnić dlaczego nie jest prawdą, że $\Delta = \Delta_+^0 - \Delta_-^0$.

•••••

Ćwiczenie. Wyjaśnić następujące zjawisko, obserwowane doświadczalnie. Jeśli w nadprzewodzącym pierścieniu płynie prąd, to strumień pola magnetycznego przez koło ograniczone przez ten pierścień może przyjmować tylko wartości

$$\Phi(B) = \frac{\pi}{e}n, \quad (2.31)$$

gdzie $(-e)$ oznacza ładunek elektronu, a n dowolną liczbę całkowitą.

Wskazówka: W nadprzewodniku nośnikami prądu są pary Coopera o ładunku $-2e$.

•••••

2.4 Równanie Schrödingera dla cząstki oddziałującej z polem elektromagnetycznym

Zależne od czasu równanie Schrödingera dla cząstki swobodnej można uogólnić dla cząstki oddziałującej z polem elektromagnetycznym, zastępując pochodne zwykle przez kowariantne i wstawiając niezależny od wycechowania potencjał $V(x)$. Otrzymujemy

$$iD^0\psi(x) = -\frac{1}{2m}\mathbf{D}^2\psi(x) + V(x)\psi(x). \quad (2.32)$$

Potencjał $V(x)$ można, na przykład, zbudować z pól \mathbf{E} i \mathbf{B} , które nie zależą od wycechowania.

Dla innego wycechowania to równanie przybiera postać

$$iD'^0\psi'(x) = -\frac{1}{2m}\mathbf{D}'^2\psi'(x) + V(x)\psi'(x), \quad (2.33)$$

ale na mocy wzoru (2.16), $D'^0\psi'(x) = (D^0\psi(x))'$ i $\mathbf{D}'^2\psi'(x) = \mathbf{D}'(\mathbf{D}\psi(x))' = (\mathbf{D}^2\psi(x))'$, czyli to równanie różni się od poprzedniego tylko dopisanym do każdego członu z lewej strony czynnikiem $e^{-iq\Lambda}$. Oba równania są więc równoważne. W ten sposób udowodniliśmy kowariancję równania (2.32) względem transformacji cechowania. Przy transformacjach Lorentza lewa strona równania transformuje się inaczej niż prawa. Nie jest to zatem równanie relatywistycznie kowariantne. Mimo to, w sytuacjach nierelatywistycznych, równanie (2.32) jest często i z powodzeniem używane.

Dla $\mathbf{A} = \mathbf{0}$, czyli w braku pola magnetycznego, można to równanie zapisać w postaci

$$i\frac{\partial\psi(x)}{\partial t} = -\frac{1}{2m}\nabla^2\psi(x) + (V(x) + qA^0(x))\psi(x), \quad (2.34)$$

a więc w postaci zwykłego równania Schrödingera z potencjałem $V(x) + qA^0(x)$. Chcąc otrzymać równanie niezależne od czasu, trzeba wstawić energię E zamiast $i\partial_t$, a nie zamiast iD_0 , jak mogło by się wydawać. Chodzi o to, żeby stan stacjonarny miał zależność od czasu daną przez czynnik e^{-iEt} . Uzasadnienie tego warunku jest przedyskutowane w paragrafie 3.1. Zgodnie z uwagą po wzorze (2.15), jeśli w definicji pochodnej kowariantnej wstawimy zamiast ładunku cząstki q dowolny stały współczynnik c , to tylko dla $c = q$ otrzymujemy klasyczny wzór na wkład potencjału elektrostatycznego do energii cząstki, a zatem tylko dla tego wyboru można zidentyfikować pole $A^\mu(x)$ ze znanym z klasycznej elektrodynamiki potencjałem wektorowym.

Sprzężenie cząstki z polem elektromagnetycznym uzyskane przez zamianę w równaniu dla cząstki swobodnej pochodnych zwykłych na kowariantne, bez dopisywania członu z $V(x)$, nazywa się **sprzężeniem minimalnym**. Przejście od równania dla cząstki swobodnej do równania ze sprzężeniem minimalnym jest tak proste, że przy omawianiu równań relatywistycznych będziemy się koncentrować na równaniach dla cząstek swobodnych, rozumiejąc że w każdej chwili można wprowadzić sprzężenie minimalne z polem elektromagnetycznym. Osobnym pytaniem jest, czy sprzężenie minimalne jest właściwe, czy trzeba dopisywać jakieś dalsze człony. O tym rozstrzyga porównanie z doświadczeniem. Wrócimy do tego zagadnienia dalej w tym paragrafie, przy dyskusji momentów magnetycznych cząstek.

Sprawdzimy teraz, jak według nierelatywistycznego równania Schrödingera elektron, mający kręt orbitalny L , oddziałuje ze stałym, jednorodnym polem magnetycznym \mathbf{B} . Moment magnetyczny $\boldsymbol{\mu}$ daje do hamiltonianu (energii) wkład

$$H'_B = -\boldsymbol{\mu}\mathbf{B}. \quad (2.35)$$

Jeśli elektron krąży, to powinien generować **moment magnetyczny**. Taki moment nazywamy orbitalnym. Energia jest skalarem względem obrotów i inwersji przestrzennej, pole magnetyczne jest pseudowektorem względem tych transformacji, więc moment magnetyczny też musi być pseudowektorem. Jedynym, poza polem magnetycznym, pseudowektorem, który da się tu utworzyć, jest kręt orbitalny L . Przewidujemy więc, że orbitalny moment magnetyczny będzie

entriw

enemag

proporcjonalny do \mathbf{L} . Znalezienie współczynnika proporcjonalności wymaga rachunku.

Wybieramy oś z wzdłuż wektora pola magnetycznego, więc $\mathbf{B} = \{0, 0, B\}$. Ze wzoru (2.20) otrzymujemy:

$$A = \left\{0, -\frac{yB}{2}, \frac{xB}{2}, 0\right\}. \quad (2.36)$$

Inne wybory potencjału wektorowego są oczywiście możliwe, ale okazują się mniej wygodne. W hamiltonianie pole magnetyczne występuje tylko w członie⁴

$$-\frac{(\nabla - iq\mathbf{A})^2}{2m} = -\frac{\nabla^2}{2m} + \frac{iq}{m}\mathbf{A}\nabla + \frac{q^2}{2m}\mathbf{A}^2. \quad (2.37)$$

Po prawej stronie, pierwszy człon nie zależy od pola magnetycznego, trzeci jest proporcjonalny do \mathbf{B}^2 , więc interesujący nas człon liniowy w \mathbf{B} jest zawarty w wyrażeniu drugim, który podobnie jak w ogólniejszym wzorze (2.35) oznaczmy H'_B . Obliczając iloczyn skalarny otrzymujemy

$$H'_B = \frac{iqB}{2m}\left(-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\right) = \frac{qB}{2m}(yp_x - xp_y) = -\frac{q}{2m}\mathbf{BL}. \quad (2.38)$$

W ostatniej równości wykorzystaliśmy fakt, że ponieważ pole magnetyczne jest równoległe do osi z , $BL_z = \mathbf{BL}$. Magneton Bohra zdefiniowany jest wzorem

$$\mu_B = \frac{e}{2m} \approx 5,8 \times 10^{-5} eV/T, \quad (2.39)$$

gdzie $(-e)$ i m oznaczają odpowiednio ładunek i masę elektronu. Dla elektronu możemy więc zapisać

$$H'_B = \mu_B \mathbf{BL}. \quad (2.40)$$

Pokazaliśmy, że ruch elektronu po orbicie generuje moment magnetyczny

$$\boldsymbol{\mu} = -\mu_B \mathbf{L}. \quad (2.41)$$

W fizyce jądrowej używa się magnetonu jądrowego:

$$\mu_N = \frac{e}{2m_p} \approx 3.15 \times 10^{-8} eV/T, \quad (2.42)$$

gdzie m_p jest masą protonu.

Obliczone powyżej oddziaływanie elektronu z polem magnetycznym okazało się niewystarczające do opisu efektu Zeemana, to znaczy rozszczepienia linii widm atomowych pod wpływem jednorodnego, stałego w czasie, pola magnetycznego. Trzeba więc dopisać do hamiltonianu jakiś człon nie wynikający ze sprzężenia minimalnego. Pauli wprowadził człon, znany obecnie jako **człon Pauli'ego**, który można zapisać w postaci

$$H''_B = g\mu_B \mathbf{SB}, \quad g = 2, \quad (2.43)$$

gdzie \mathbf{S} oznacza spin elektronu. Moment magnetyczny, który można odczytać z tego wzoru, jest własnością elektronu i nie ma nic wspólnego z ruchem orbitalnym. Współczynnik g jest znany jako współczynnik żyromagnetyczny

⁴W naszym przykładzie $\nabla \cdot \mathbf{A} = 0$, więc ∇ i \mathbf{A} komutują.

i został pierwotnie wyznaczony z doświadczenia. Fakt, że ten współczynnik jest różny od jedności komplikuje opis efektu Zeemana. Stany własne energii atomów, kiedy nie ma pól zewnętrznych, wybiera się zwykle tak, że mają dobrze określoną liczbę kwantową $J_z = L_z + S_z$, a energia magnetyczna jest proporcjonalna do $L_z + 2S_z$. Równanie Schrödingera z członem Pauli'ego nieźle opisuje efekt Zeemana, w atomie wodoru i atomach metali alkalicznych, w silnych polach magnetycznych, ale przy słabych polach magnetycznych nabierają znaczenia dodatkowe oddziaływania, które omówimy przy dyskusji równania Diraca. Zauważmy, że spin jest jedynym pseudowektorem wchodzącym do opisu cząstki w spoczynku. Wobec tego moment magnetyczny takiej cząstki musi być proporcjonalny do \mathbf{S} . Zmierzyć, lub wyliczyć, trzeba tylko współczynnik żyromagnetyczny.

W rozdziale V pokażemy, że człon Pauli'ego ze współczynnikiem żyromagnetycznym $g = 2$ wynika z równania Diraca bez wprowadzania żadnych członów poza sprzężeniem minimalnym. Dalsze prace doświadczalne wykazały jednak, że współczynnik żyromagnetyczny odbiega nieco od wartości dwa otrzymanej przez Diraca. To odchylenie wyjaśnia elektrodynamika kwantowa, która daje

$$\frac{g-2}{2} = \frac{\alpha}{2\pi} + O(\alpha^2). \quad (2.44)$$

Ponieważ $\alpha \approx \frac{1}{137}$, poprawka jest bardzo mała. Została natomiast zmierzona z ogromną dokładnością. Błąd doświadczalny wynosi obecnie [15] 2.7×10^{-13} . Rachunek w ramach elektrodynamiki kwantowej, uwzględniający człony z wyższymi potęgami α i bardzo niewielką poprawkę hadronową, daje wynik zgodny z doświadczeniem w granicach tego błędu. Podobnie wygląda sytuacja dla mionu. Natomiast dla nukleonów równanie Diraca nie pozwala dobrze oszacować członu Pauli'ego, przy czym nie chodzi tu tylko o modyfikację współczynnika żyromagnetycznego. Neutron jest obojętny elektrycznie, więc powinien mieć zerowy moment magnetyczny. Tymczasem doświadczalnie moment magnetyczny neutronu wynosi około $-1,91\mu_N$, i co do wartości bezwzględnej jest porównywalny z momentem magnetycznym protonu (około $2.79\mu_N$), który zresztą też się nie zgadza z przewidywaniem teorii Diraca. Można to zrozumieć, jeśli się wie, że wprawdzie neutron jako całość jest elektrycznie obojętny, ale składa się z cząstek o niezerowych ładunkach elektrycznych. Z faktu, że ich ładunki elektryczne się znoszą, nie wynika, że znoszą się także ich momenty magnetyczne.

Rozdział 3

Równanie Kleina-Gordona

3.1 Równanie Kleina-Gordona i jego rozwiązania dla cząstki swobodnej

Równanie Kleina-Gordona jest podstawowym relatywistycznym równaniem dla funkcji falowych¹ cząstek o spinie zero. Funkcje falowe cząstek o wyższych spinach również spełniają to równanie, ale spełniają też inne równania, które dają więcej informacji.

Równanie Kleina-Gordona dla cząstki swobodnej ma postać

$$(\hat{p}_\mu \hat{p}^\mu - m^2)\psi(x) = 0. \quad (3.1)$$

Jak widać, powstaje ono z klasycznego wzoru $p^2 - m^2 = 0$, podobnie jak niere-relatywistyczne równanie Schrödingera dla cząstki swobodnej powstaje ze wzoru $\frac{\mathbf{p}^2}{2m} = E$. Podstawiając $p^\mu = i\partial^\mu$ i zmieniając znak lewej strony równania otrzymujemy

$$(\partial_\mu \partial^\mu + m^2)\psi(x) = 0. \quad (3.2)$$

Zwykle zapisuje się to równanie przy pomocy d'Alambertjana, $\square \equiv \partial_\mu \partial^\mu$: *klegor*

$$(\square + m^2)\psi(x) = 0. \quad (3.3)$$

Równanie jest liniowe, więc jego rozwiązania stanowią przestrzeń wektorową.

Omówimy teraz działanie operatorów translacji. Ten przypadek jest tak prosty, że można by od razu napisać końcowe wyniki. Ogólna dyskusja przenosi się jednak z bardzo niewielkimi zmianami na bardziej skomplikowane operatory, a łatwiej jest ją zrozumieć na tym prostym przykładzie. Niech operator translacji $\hat{T}(a)$ zmienia każdą punktochwilę x w $x' = x + a$: *trapch*

$$\hat{T}(a)x = x' = x + a, \quad x = \hat{T}^{-1}(a)x' = x' - a. \quad (3.4)$$

Wprowadzamy oznaczenie

¹Zgodnie ze zwyczajem, rozwiązania $\psi(x)$ relatywistycznych równań falowych będziemy nazywali funkcjami falowymi, chociaż z kwantowej teorii pola wiadomo, że to są klasyczne pola, które nie dadzą się interpretować jako iloczyny skalarne $\langle x|\psi\rangle$, gdzie $|\psi\rangle$ jest wektorem stanu. W związku z tym nie będziemy używać notacji Diraca dla tych funkcji.

$$f'(x) = \hat{T}(a)f(x). \quad (3.5)$$

fprxpr

Dla translacji przyjmujemy, że

$$f'(x') = f(x). \quad (3.6)$$

fprtra

Interpretację fizyczną tego założenia dyskutujemy poniżej. Korzystając z drugiego ze wzorów (3.4), możemy wyeliminować argument x . Zmieniając następnie oznaczenie z x' na x otrzymujemy

$$f'(x) = f(\hat{T}^{-1}(a)x) \quad (3.7)$$

Ćwiczenie Wyjaśnić dlaczego ten wzór nie jest sprzeczny ze wzorami (3.4) dla $f(x) = x$.

•••••

Powyższy wzór wyjaśnia sens fizyczny założenia (3.6). Wartość funkcji f' w punktochwili x jest równa wartości funkcji f w punktochwili x , która pod wpływem translacji o a przejdzie w punktochwile x . Ogólnie, jeśli wzór (3.6) jest prawdziwy dla wszystkich operatorów z jakiejś grupy, to mówimy, że funkcja f jest skalarem względem tej grupy. W takim przypadku prawdziwy jest wzór analogiczny do wzoru (3.7), bo w jego wyprowadzeniu nie korzystaliśmy z żadnych szczególnych własności operatorów translacji. Jeżeli funkcja nie jest skalarna, to na ogół wzór (3.7) się nie stosuje. Na przykład, jednorodne pole elektryczne równoległe do osi z przechodzi przy obrocie o 90° wokół osi y w jednorodne pole elektryczne równoległe do osi x . Dotyczy to wszystkich punktów, nawet punktów na osi y , które przy tym obrocie nie zmieniają położenia ($\mathbf{x}' = \mathbf{x}$). Operatory translacji zawsze można zdefiniować tak, że wzór (3.6) stosuje się.

Ćwiczenie Na przykładzie pola temperatur na płaszczyźnie, $T(x, y)$, które jest skalarnie względem obrotów wokół osi z , wyjaśnić sens fizyczny wzoru (3.6). Jako transformację przyjąć obrót wokół osi z o kąt ϕ .

•••••

fprima

W ogólnym przypadku, kiedy jakiś operator \hat{A} działa na funkcję $f(x)$, która nie musi być skalarna, odpowiednikiem wzoru (3.7) jest wzór

$$f'(x) = \tilde{f}(\hat{A}^{-1}x), \quad (3.8)$$

gdzie funkcję \tilde{f} można znaleźć rozpatrując szczególny przypadek, kiedy funkcja $f(x)$ nie zależy od punktochwili x . Jeżeli

$$\hat{A}^{-1} = \hat{A}, \quad (3.9)$$

fprimb

to można używać nieco prostszego wzoru

$$f'(x) = \tilde{f}(\hat{A}x), \quad (3.10)$$

Ćwiczenie Pokazać, że dla pola wektorowego $f(\mathbf{r}) = \mathbf{r} \cos \phi$ wzór (3.7) stosuje się dla translacji, ale nie dla rotacji wokół osi z . Znaleźć uogólnienie, które stosuje się do tych rotacji i porównać ze wzorem (3.8).

•••••

Wracamy do równania Kleina-Gordona. Ponieważ różniczkowanie po x daje ten sam wynik, co różniczkowanie po $x + a$, mamy:

$$0 = \hat{T}(a)(\square + m^2)\psi(x) = (\square + m^2)\hat{T}(a)\psi(x). \quad (3.11)$$

To znaczy, że w przestrzeni rozwiązań równania Kleina-Gordona operatory $\hat{T}(a)$ i $\square + m^2$ komutują ze sobą. Można więc znaleźć układ zupełny rozwiązań równania Kleina-Gordona wśród funkcji własnych operatorów $\hat{T}(a)$. Te funkcje mają postać

falpla

$$\psi_p(x) = N(p)e^{-ipx}, \quad (3.12)$$

gdzie współczynnik $N(p)$ nie zależy od x , ale może zależeć od p . Rzeczywiście:

$$\hat{T}(a)N(p)e^{-ipx} = N(p)e^{-ip(x-a)} = e^{ipa}N(p)e^{-ipx}, \quad (3.13)$$

Funkcja falowa (3.12) jest też wspólną funkcją własną operatorów składowych czteropędu, $\hat{p}^\mu = i\partial^\mu$, do wartości własnych p^μ . To znaczy, że składowe czteropędu cząstki, której stan opisuje ta funkcja falowa, są dobrze określone i mają wartości p^μ . Dla jednowymiarowych funkcji falowych czynnik $N(p)$ jest po prostu stałą normalizacyjną. Dla funkcji dwu lub więcej komponentowych, ten czynnik jest bardziej skomplikowaną funkcją pędu, która pozostaje do wyznaczenia.

Ćwiczenie Rozważając translacje $\hat{T}(a)$, dla których moduły wszystkich składowych czterowektora a są bardzo małe, wyjaśnić dlaczego wspólne funkcje własne operatorów składowych czteropędu są zarazem funkcjami własnymi operatorów translacji.

• • • • •

Zauważmy, że powyższe wyniki można uzyskać zakładając tylko, że żaden punkt, czy obszar, w czasoprzestrzeni nie jest wyróżniony, bo to znaczy, że operator działający na funkcje falową, w naszym przypadku operator $\square + m^2$ jest niezmienniczy względem translacji. To założenie jest spełnione przez wszystkie równania dla cząstek swobodnych. Dlatego każde z tych równań będzie miało rozwiązania postaci (3.12). Ale, jak zaraz zobaczymy na przykładzie równania Kleina-Gordona, nie wszystkie te funkcje są rozwiązaniami równania. Na ogół równanie nakłada jeszcze dodatkowe warunki.

Jeżeli w równaniu występuje oddziaływanie niezależne od czasu, to można powtórzyć powyższe rozumowanie ograniczając się do przesunięć w czasie. Odpowiednikiem wzoru (3.12) jest wtenczas wzór

faltem

$$\psi_{p^0}(x) = e^{-ip^0 t} \psi_{p^0}(0, \mathbf{x}). \quad (3.14)$$

Podstawiając funkcje $\psi_p(x)$ do równania (3.3), otrzymujemy:

$$(p^0)^2 - \mathbf{p}^2 = m^2, \quad (3.15)$$

lub równoważnie

$$p^0 = \pm \sqrt{\mathbf{p}^2 + m^2}. \quad (3.16)$$

Rozwiązania z $p^0 > 0$ opisują cząstki z pędem \mathbf{p} i energią $E = p^0$. Rozwiązania z $p^0 < 0$ stanowiły problem, bo w doświadczeniu nie obserwuje się cząstek swobodnych z ujemną energią.

Gdybyśmy chcieli opisywać tylko cząstki swobodne, moglibyśmy się umówić, że dopuszczamy tylko stany z dodatnią energią, ale takie podejście jest nie do utrzymania, jeśli są oddziaływania z polami zewnętrznymi². Wtedy, nawet jeśli dla $t = 0$ funkcja falowa jest superpozycją tylko funkcji z dodatnimi energiami, pojawiają się w miarę ewolucji układu składowe z energiami ujemnymi. Na skutek tego, całkowite prawdopodobieństwo, że cząstka ma energię dodatnią, spada poniżej jedności i konieczna jest jakaś interpretacja pozostałych możliwości związanych ze składowymi funkcji falowej z ujemnymi energiami. Jeżeli uznamy, że poziomy z ujemną energią są fizyczne, to mamy kolejną trudność. Cząstki z dodatnią energią będą spadać na te poziomy obniżając nieograniczenie swoje energie.

Pierwsze rozwiązanie tej trudności podał Dirac, który natknął się na nią przy analizie swojego równania. Założył, że w stanie podstawowym układu wszystkie stany o energii ujemnej są obsadzone. Ponieważ rozważał elektrony, mógł się powołać na zakaz Pauli'ego, który uniemożliwia cząstkom z energią dodatnią spадanie na już zajęte poziomy z energią ujemną. Ten zbiór elektronów o ujemnych energiach nazywa się morzem Diraca. Z interpretacji Diraca wynika ważny wniosek. Może się zdarzyć, że elektron z ujemną energią pochłonie kwant promieniowania elektromagnetycznego o tak dużej energii, że stanie się elektronem z dodatnią energią. Mamy wtedy produkcję pary: pojawia się elektron o energii dodatniej i pozostała po nim dziura w morzu Diraca. Energię morza liczymy jako sumę energii składających się na nie elektronów, bez uwzględniania oddziaływań międzyelektronowych. Energię i pęd dziury obliczamy przez odjęcie czteropędu morza Diraca z dziurą od jego czteropędu bez tej dziury. Nie przejmujemy się tym, że odejmujemy od siebie dwie wielkości nieskończone. Stąd, jeżeli elektron miał energię $p^0 < 0$ i pęd \mathbf{p} , to dziura ma energię $E = -p^0 > 0$ i pęd $-\mathbf{p}$. Ma więc dodatnią energię, i masę $\sqrt{E^2 - \mathbf{p}^2}$ taką samą jak elektron. Podobnie, rzut spinu S_z na dowolnie wybraną oś z jest dla dziury przeciwny niż dla odpowiadającej jej cząstki z ujemną energią. Wynika stąd, że spin S dziury jest taki sam jak spin odpowiadającej jej cząstki. Zauważmy jeszcze, że skrętność nie ulega zmianie, bo zmieniają się jednocześnie zwroty wektorów pędu i spinu. W polu elektrycznym dziura porusza się w kierunku przeciwnym do kierunku ruchu elektronów, podobnie jak pęcherzyk powietrza w wodzie porusza się ku górze, a więc w kierunku przeciwnym do kierunku siły grawitacji działającej na cząsteczki wody. Wnioskujemy stąd, że dziura zachowuje się jak cząstka o ładunku przeciwnym do ładunku elektronu. Tę cząstkę nazywamy antycząstką. Z teorii Diraca wynika więc, że dla elektronu, ogólniej dla każdego fermionu, istnieje jego antycząstka, czyli cząstka o takiej samej masie i spinie, ale przeciwnym ładunku elektrycznym. Doświadczalne potwierdzenie tego przewidywania dla elektronu było wielkim sukcesem teorii Diraca. Teorię równoważną z teorią Diraca można zbudować przyjmując, że to pozytony są cząstkami z dodatnią energią, a elektrony są dziurami w morzu Diraca pozytonów. To, że elektrony nazywamy cząstkami, a pozytony antycząstkami ma uzasadnienie wyłącznie historyczne.

Interpretacja Diraca, odpowiednio zmodyfikowana, odgrywa ważną rolę w fizyce fazy skondensowanej, gdzie nośnikami prądu elektrycznego mogą być zarówno ujemne elektrony jak i dodatnie dziury. Należy jednak pamiętać, że tam

²Przyjmujemy tu bez dowodu dość oczywisty fakt, że wprowadzenie oddziaływań, na ogół, nie eliminuje rozwiązań z ujemną energią

zamiast morza Diraca mamy morze Fermi’ego złożone z elektronów oddziałujących z jonami. Na skutek tego masa dziury jest na ogół różna od masy elektronu i obie są na ogół różne od masy elektronu figurującej w tablicach cząstek. W fizyce cząstek interpretacja Diraca ma dużą wartość heurystyczną — zastanawiając się nad jakąś własnością antycząstek, często pomocne jest sprawdzenie, jak to jest przy interpretacji Diraca. Z doświadczenia wiadomo jednak, że nie tylko fermiony mają antycząstki. Mają je także bozony, dla których nie obowiązuje zakaz Pauli’ego, a więc interpretacja Diraca nie da się zastosować. Interpretację dającą się zastosować zarówno do fermionów jak i do bozonów podał Stückelberg (1941), a następnie znakomicie rozwinął Feynman (1948). Dodatkową zaletą tej nowszej interpretacji jest, że nie wprowadza nieskończonej liczby cząstek o nieskończonej (ujemnej) łącznej energii.

Według Stückelberga i Feynmana, antycząstka odpowiada cząstce o ujemnej energii poruszającej się w tył w czasie. Z tego określenia natychmiast wynika, że ma taką samą masę i taki sam spin jak cząstka. Człon $-ip^0\Delta t$, występujący w czynniku fazowym funkcji falowej stanu stacjonarnego, przechodzi w człon $-ip^0(-\Delta t) = -i(-p^0)\Delta t$. Dla stanu o ujemnej energii $-p^0 = E > 0$ i otrzymujemy człon jak dla normalnej cząstki z energią $E > 0$ poruszającej się w przód w czasie. Antycząstka ma ładunek elektryczny przeciwny niż cząstka. Poruszanie się w tył w czasie w sensie Stückelberga i Feynmana należy więc odróżniać od znanego z klasycznej fizyki odwracania ruchu w czasie, które nie tworzy nowej cząstki z przeciwnym ładunkiem. Pokażemy teraz, że ta interpretacja też przewiduje powstawanie par. Rozważmy linię świata elektronu poruszającego się w tył w czasie do chwili $t = t_0$. W chwili t_0 linia świata zwraca i elektron dalej porusza się w przód w czasie. Co widzi obserwator? Dla $t < t_0$ nie ma żadnego elektronu. Dla każdego $t > t_0$ widać jeden elektron poruszający się w tył w czasie i jeden poruszający się w przód w czasie. Powstała więc para e^+e^- . Z drugiej strony, mamy tylko jedną linię świata. Feynman następująco popularyzował tę zmianę punktu widzenia. Wyobraźmy sobie, że linia świata jest drogą. Kierowca ciężarówki jadący tą drogą zawsze widzi jedną drogę. Natomiast lotnik lecący w kierunku wzrastającej współrzędnej czasowej najpierw nie widzi żadnej drogi, a potem pojawiają mu się dwie. Zauważmy, że interpretacja Stückelberga-Feynmana też jest symetryczna względem zamiany cząstek na antycząstki. Używając tej interpretacji, można opisywać niektóre efekty zwykle uważane za wielocząstkowe jako procesy jednocząstkowe. Przykładem jest produkcja par, którą zwykle uważa się za proces dwucząstkowy. Potrzebnej energii może dostarczać, na przykład, klasyczne pole elektromagnetyczne. To rozszerzenie klasy procesów, które dają się opisywać przy pomocy mechaniki kwantowej, gdzie liczba cząstek się nie zmienia, bardzo poprawia skuteczność metody zaburzeń. Dobrą dyskusję tego zagadnienia, ilustrowaną licznymi przykładami, można znaleźć w podręczniku [7]. Cząstki z ujemną energią poruszające się w tył w czasie mogą być obserwowane tylko jako antycząstki o dodatniej energii poruszające się w przód w czasie. Ich istnienie nie narusza więc zasady przyczynowości.

Ćwiczenie Uzasadnić, używając interpretacji Stückelberga-Feynmana, że antycząstka ma ładunek elektryczny przeciwny w stosunku do odpowiadającej jej cząstki.



Obie powyższe interpretacje historycznie odegrały wielką rolę i pozwoliły wy-

prorowadzić cenne wyniki, ale według obecnych poglądów w pełni konsyistentną interpretację rozwiązań z ujemną energią daje dopiero kwantowa teoria pola. Tam rozwiązania z ujemną energią są wykorzystywane, ale nie prowadzą do pojawienia się cząstek z ujemną energią. W praktyce, wyniki pokrywają się z wynikami które można uzyskać używając interpretacji Stückelberga-Feynmana.

3.2 Prąd zachowywany

W teorii nierelatywistycznej kwadrat modułu odpowiednio znormalizowanej funkcji falowej interpretujemy jako gęstość prawdopodobieństwa. Jest to możliwe tylko dlatego, że istnieje prąd J spełniający równanie ciągłości $\partial_\mu J^\mu = 0$, i że stałą normalizacyjną można wybrać tak, że $J^0 = |\psi(x)|^2$. Prawdopodobieństwo znalezienia cząstki gdziekolwiek w przestrzeni, które musi być identycznie równe jeden, jest zachowywane na mocy równania ciągłości. Jest to tylko warunek konieczny, który musi spełniać gęstość prawdopodobieństwa. Trzeba jeszcze sprawdzić, że interpretacja części przestrzennej J jako prądu gęstości prawdopodobieństwa jest rozsądna.

Ćwiczenie Pokazać, że stały w czasie i w przestrzeni wektor j^μ spełnia równanie ciągłości i wyjaśnić dlaczego, mimo to, na ogół, wielkość j^0 nie może być interpretowana jako gęstość prawdopodobieństwa.

•••••

Żeby znaleźć interpretację probabilistyczną rozwiązań równania Kleina-Gordona, musimy najpierw wykazać, że i tu istnieje prąd zachowywany, to znaczy spełniający równanie ciągłości. W tym celu mnożymy równanie lewostronnie przez $\psi^*(x)$, a następnie równanie sprzężone zespolone lewostronnie przez $\psi(x)$. Otrzymujemy układ dwu równań

$$\psi^*(x)(\square + m^2)\psi(x) = 0; \quad \psi(x)(\square + m^2)\psi^*(x) = 0. \quad (3.17)$$

Odejmując je od siebie stronami, mamy równość

$$\psi^*(x)\square\psi(x) - \psi(x)\square\psi^*(x) = 0. \quad (3.18)$$

Korzystając z tożsamości

$$0 = \psi^*(x)\partial_\mu\partial^\mu\psi(x) - \psi(x)\partial_\mu\partial^\mu\psi^*(x) = \partial_\mu [\psi^*(x)\partial^\mu\psi(x) - \psi(x)\partial^\mu\psi^*(x)] \quad (3.19)$$

otrzymujemy równanie ciągłości

$$\partial_\mu J^\mu = 0, \quad (3.20)$$

jeśli przyjmiemy

$$J_\mu = N (\psi^*\partial_\mu\psi - \psi\partial_\mu\psi^*), \quad (3.21)$$

gdzie N jest dowolną stałą. Dodanie do m skalarne go potencjału $V(x)$ nie wpływa na równanie ciągłości. Wprowadzenie sprzężonego minimalnie pola elektromagnetycznego daje, po zwykłych podstawieniach $\partial_\mu\psi \rightarrow D_\mu\psi$ i $\partial_\mu\psi^* \rightarrow D_\mu^*\psi^*$, równanie ciągłości z prądem

$$J_\mu = N (\psi^* D_\mu \psi - \psi D_\mu^* \psi^*), \quad (3.22)$$

gdzie $D_\mu = \partial_\mu + iqA_\mu$ oznacza, jak zwykle, pochodną kowariantną, a N stałą normalizacyjną.

Ćwiczenie Pokazać, że prąd (3.22) jest zachowywany i nie zależy od wycechowania.

•••••

W szczególności, dla cząstki oddziałującej z niezależnym od czasu potencjałem skalarnym, w stanie stacjonarnym zachodzi $\psi(x) = e^{-ip^0 t} \psi(0, \mathbf{x})$, i otrzymujemy *geprkg*

$$J^0(x) = -2iNp^0 |\psi(0, \mathbf{x})|^2. \quad (3.23)$$

Dla $p^0 > 0$ można wybrać $N = \frac{i}{2m}$ i otrzymać

$$J^0(x) = \frac{p^0}{m} |\psi(0, \mathbf{x})|^2. \quad (3.24)$$

To wyrażenie można interpretować jako gęstość prawdopodobieństwa. W granicy nierelatywistycznej $\frac{p^0}{m} \rightarrow 1$, i odtwarzamy znany wzór nierelatywistyczny.

Natomiast dla $p^0 < 0$ nie możemy już inaczej wybrać współczynnika $-iN$ i interpretacja probabilistyczna nie ma sensu, bo $J^0(x) \leq 0$. To był kiedyś poważny argument przeciwko używaniu równania Kleina-Gordona.

Ćwiczenie Przedyskutować przypadek, kiedy funkcja falowa jest superpozycją rozwiązań z dodatnimi i ujemnymi energiami i pokazać, że dla prądu zachowywanego nie ma możliwości wprowadzenia współczynnika N zależnego od znaku p_0 .

•••••

Wyjście znaleźli Pauli i Weisskopf (1934), którzy zaproponowali, żeby funkcję

$$J^0(x) = q \frac{E}{m} |\psi(0, \mathbf{x})|^2, \quad (3.25)$$

gdzie q oznacza elektryczny ładunek cząstki, a $E = |p^0|$, interpretować jako gęstość ładunku elektrycznego. Ponieważ cząstki i antycząstki mają przeciwne ładunki, związane z nimi gęstości ładunku muszą mieć przeciwne znaki, zgodnie z powyższym równaniem. Na mocy równania ciągłości zachowywany jest więc całkowity ładunek cząstek, a nie ich liczba, co jest zgodne z doświadczeniem. Zauważmy jeszcze, że przy tej interpretacji dla cząstek elektrycznie neutralnych $q = 0$, więc $J^\mu(x) \equiv 0$. W tym przypadku nie ma zachowywanego prądu i interpretację probabilistyczną można podać dopiero na gruncie kwantowej teorii pola.

Czynnik $\frac{E}{m} = \gamma$ poprawia na skrócenie Lorentza. Przy zmniejszeniu elementu objętości o czynnik γ (skrócenie Lorentza), bez zmiany liczby cząstek w tym elemencie, gęstość wzrasta o czynnik γ .

3.3 Symetrie równania Kleina-Gordona

Operacja A jest symetrią równania falowego

rowfal

$$\hat{O}\psi(x) = 0, \quad (3.26)$$

jeśli dla każdego rozwiązania $\psi(x)$, funkcja

$$\psi^A(x) \equiv \hat{A}\psi(x). \quad (3.27)$$

też jest rozwiązaniem. Wtedy

$$\hat{A}\hat{O}\psi(x) = 0 = \hat{O}\hat{A}\psi(x). \quad (3.28)$$

Ponieważ rozwiązania $\psi(x)$ stanowią układ zupełny, w przestrzeni rozwiązań równania (3.26), wynika stąd równoważne określenie symetrii: operacja A jest symetrią równania (3.26) kiedy komutator

$$[\hat{A}, \hat{O}] = 0, \quad (3.29)$$

czyli kiedy operator \hat{O} nie zmienia się pod działaniem operatora \hat{A} . W paragrafie 3.1 omawialiśmy już szczególny przypadek, kiedy $\hat{O} = \square + m^2$ i $\hat{A} = \hat{T}(a)$.

Przedstawimy teraz najważniejsze symetrie równania Kleina-Gordona cząstki swobodnej i przedyskutujemy ich stosowalność dla innych równań.

Właściwe transformacje Lorentza Operatory $L_{\uparrow+}$ komutują ze skalarem $\square = \partial_\mu \partial^\mu$, więc, jeżeli funkcja $\psi(x)$ jest rozwiązaniem swobodnego równania Kleina-Gordona, to zachodzi też

$$(\square + m^2)\hat{L}_{\uparrow+}\psi(x) = 0. \quad (3.30)$$

Funkcja $\hat{L}_{\uparrow+}\psi(x)$ może być interpretowana jako funkcja falowa $\psi(x)$ widziana z innego układu odniesienia. Z zasady względności wynika więc, że równanie falowe dla tej funkcji musi być takie same jak dla funkcji $\psi(x)$. Zasada względności obowiązuje dla każdego układu fizycznego, więc właściwe transformacje Lorentza są symetrami wszelkich realistycznych, relatywistycznych równań dla cząstek³.

Inwersja przestrzenna Operator inwersji przestrzennej, krócej inwersji lub parzystości, oznaczamy symbolem \hat{P} . Pod działaniem inwersji

$$\hat{P}x = \{t, -\mathbf{x}\}, \quad \hat{P}^2x = x. \quad (3.31)$$

Z drugiego z tych wzorów nie wynika, że $\hat{P}^2 = 1$, ale tylko że **reprezentant** tego operatora działający na punktochwile ma tę własność. Może się wydawać oczywiste, że dwukrotna inwersja przestrzenna nic nie zmienia, ale równie dobry argument, że obrót o kąt 2π nic nie zmienia jest, jak wiadomo, fałszywy dla funkcji falowych cząstek o spinie połówkowym, czyli dla fermionów. Przyjmijmy więc tylko że

$$\hat{P}^2 = 1 \quad (3.32)$$

w działaniu na funkcje falowe cząstek o spinach całkowitych, czyli dla bozonów. Fermionami zajmiemy się w następnym rozdziale.

Ograniczymy się teraz do cząstek o spinie zero. Ich funkcje falowe będziemy oznaczali $\phi(x)$. Dla takich cząstek funkcja falowa jest jedną liczbą i funkcje ϕ i $\bar{\phi}$ mogą się różnić co najwyżej o stały czynnik fazowy, który tu oznaczymy ϵ . Mamy więc

³Jak wspominaliśmy w rozdziale pierwszym może tak nie być w silnych polach grawitacyjnych

$$\hat{P}\phi(x) = \epsilon\phi(\hat{P}x), \quad \hat{P}^2\phi(x) = \epsilon^2\phi(x) \quad (3.33)$$

Ponieważ dla cząstki o spinie zero $\hat{P}^2 = 1$, musi zachodzić $\epsilon^2 = 1$, a zatem

$$\epsilon = \pm 1. \quad (3.34)$$

Znak w tym wzorze zależy od rodzaju opisywanej cząstki i nazywa się jej **parzystością wewnętrzną**. Cząstki o spinie zero i parzystości wewnętrznej plus są nazywane skalarnymi. Przykładem jest bozon Higgsa. Cząstki o spinie zero i parzystości wewnętrznej minus, są nazywane pseudoskalarnymi. Przykładem jest mezon π . Analogiczne wyniki uzyskuje się dla bozonów o wyższych spinach całkowitych S . Mówimy, że parzystość bozonu jest naturalna, jeśli $\epsilon = (-1)^S$, lub nienaturalna, jeśli $\epsilon = (-1)^{S+1}$. Podobnie jak dla $S = 0$, w określeniu typu cząstki o parzystości nienaturalnej dodajemy przedrostek *pseudo*. Tak więc na przykład, dla $S = 1$ cząstki z ujemną parzystością wewnętrzną nazywamy wektorowymi, a cząstki z dodatnią parzystością wewnętrzną pseudowektorowymi. Wszystkie znane bozony mają dobrze określone parzystości wewnętrzne.

Może się zdarzyć, że funkcja falowa cząstki spełnia relację

$$\phi(\hat{P}x) = \epsilon_L\phi(x) \quad (3.35)$$

Czynnik fazowy $\epsilon_L = \pm 1$ nazywamy **parzystością orbitalną**. W takim przypadku

$$\hat{P}\phi(x) = \epsilon\phi(\hat{P}x) = \epsilon\epsilon_L\phi(x). \quad (3.36)$$

Funkcja $\phi(x)$ jest więc funkcją własną operatora parzystości. Wartość własną $\epsilon\epsilon_L$ nazywa się krótko parzystością i to ta **parzystość** jest zachowywana, jeśli \hat{P} jest symetrią równania. Przy większej niż jeden liczbie cząstek, parzystość jest iloczynem parzystości wewnętrznych wszystkich cząstek i ich wspólnej parzystości orbitalnej.

Ćwiczenie Sprawdzić, że dla atomu wodoru w stanie o kręcie orbitalnym L , parzystość orbitalna $\epsilon_L = (-1)^L$.

•••••

Ćwiczenie Przedyskutować jakościowo inwersję stanu cząstki skalarniej, z dodatnim ładunkiem elektrycznym, umieszczonej między okładkami kondensatora. Rozpatrzeć przypadki: kiedy inwersja obejmuje okładki kondensatora i kiedy ich nie obejmuje.

•••••

Ćwiczenie Wyjaśnić dlaczego w nierelatywistycznej mechanice kwantowej, gdzie żadna cząstka nigdy nie zmienia się w inną, nie wprowadza się parzystości wewnętrznych.

•••••

Ponieważ zmiana znaku \mathbf{x} nie wpływa na operator $\partial_\mu\partial^\mu$, \hat{P} jest operatorem symetrii równania Kleina-Gordona dla cząstki swobodnej. W danym polu zewnętrznym, funkcja $\phi^P(x)$ nie spełnia na ogół równania Kleina-Gordona, które spełniała funkcja $\phi(x)$. Nie ma więc symetrii. Symetria w obecności oddziaływań dotyczy tylko sytuacji, kiedy inwersji podlega nie tylko funkcja falowa

cząstki, ale i pola, czy równoważnie ich źródła, z którymi cząstka oddziałuje. W tym sensie parzystość jest zachowywana (jest dobrą symetrią) w oddziaływaniach silnych i elektromagnetycznych, jest natomiast łamana w oddziaływaniach słabych.

Ćwiczenie Pokazać, że jeśli cząstka rozpada się przez oddziaływania silne lub elektromagnetyczne na dwa mezony π , to musi mieć parzystość naturalną.

•••••

Inwersja przestrzenna może być przedstawiona jako iloczyn dwu operacji: zwierciadlanego odbicia w płaszczyźnie x, y i obrotu wokół osi z o kąt 180° . Ponieważ obroty wokół osi z stanowią podgrupę grupy $\hat{L}_{\uparrow+}$, a zatem są dobrą symetrią, pytanie czy inwersja jest dobrą symetrią jest równoważne z pytaniem, czy świat widziany po drugiej stronie lustra podlega tym samym prawom, co ten po naszej. Żeby się przekonać, że dla słabych oddziaływań tak nie jest, rozważmy produkcję neutrina. Ta produkcja zachodzi wyłącznie poprzez słabe oddziaływania. W doświadczeniu po naszej stronie lustra obserwuje się tylko neutrina lewoskrętne. Po drugiej stronie lustra wszystkie te neutrina stają się prawoskrętne, a więc świat po tamtej stronie jest rządony innymi prawami – inwersja przestrzenna nie jest dobrą symetrią dla słabych oddziaływań, parzystość nie jest zachowywana.

Sprzężenie ładunkowe Wprowadzamy operator sprzężenia ładunkowego \hat{C} , który nie zmienia punktochwil, natomiast powoduje sprzężenie zespolone funkcji falowej. Mamy więc

$$\phi^C(x) = \epsilon_C \phi^*(x). \quad (3.37)$$

Parametr ϵ_C nazywa się parzystością ładunkową cząstki. Ze wzoru widać, że mnożąc funkcję $\phi(x)$ przez stały czynnik fazowy można by zmienić parametr ϵ_C . Ponieważ operator sprzężenia ładunkowego jest antyliniowy,

$$\hat{C}^2 \phi(x) = |\epsilon_C|^2 \phi(x), \quad (3.38)$$

więc warunek $\hat{C}^2 = 1$ nie daje żadnej informacji o współczynniku fazowym ϵ_C . Biorąc sprzężenie zespolone równania Kleina-Gordona dla cząstki swobodnej widać, że sprzężenie ładunkowe jest symetrią tego równania.

Równanie Kleina-Gordona dla cząstki o ładunku q , sprzężonej z polem elektromagnetycznym,

$$(\partial_\mu + iqA_\mu(x))(\partial^\mu + iqA^\mu(x))\phi(x) + m^2\phi(x) = 0, \quad (3.39)$$

pod działaniem sprzężenia zespolonego, po pomnożeniu przez ϵ_C , przybiera postać

$$(\partial_\mu - iqA_\mu(x))(\partial^\mu - iqA^\mu(x))\phi^C(x) + m^2\phi^C(x) = 0. \quad (3.40)$$

To jest równanie dla funkcji falowej cząstki o tej samej masie i spinie, ale o ładunku $-q$, czyli dla antycząstki. Dlatego operator \hat{C} nazywa się operatorem sprzężenia ładunkowego. Zauważmy jeszcze, że pod wpływem operacji \hat{C} czynnik e^{-ip_0t} przechodzi w e^{+ip_0t} , a więc w szczególności funkcje falowe cząstek z $p_0 < 0$, czyli z ujemną energią, przechodzą w funkcje falowe antycząstek z energią dodatnią.

Podobnie jak parzystość, sprzężenie ładunkowe nie jest dobrą symetrią, kiedy dotyczy tylko oddziałującej cząstki. Kiedy jednak sprzężenie ładunkowe obejmuje także źródła pól zewnętrznych, mamy transformację $qA \rightarrow (-q)(-A) = qA$, i równanie Kleina-Gordona pozostaje bez zmian. Udowodniliśmy więc, że przy tej interpretacji sprzężenie ładunkowe jest dobrą symetrią dla równania Kleina-Gordona z minimalnym sprzężeniem elektromagnetycznym. Ogólniej okazuje się, że jest to dobra symetria dla oddziaływań elektromagnetycznych (nie koniecznie ze sprzężeniem minimalnym) i silnych. Nie jest natomiast symetrią oddziaływań słabych. Kontrprzykładu dostarcza znów produkcja neutrin: sprzężenie ładunkowe zamienia lewoskrętne neutrino w nieobserwowane doświadczalnie lewoskrętne antyneutrino. Zauważmy jeszcze, że ponieważ $(\phi^*(x))^* = \phi(x)$, mamy do wyboru, którą z cząstek opisywanych funkcjami falowymi $\phi(x)$ i $\phi^C(x)$ uznać za cząstkę, a którą za antycząstkę. Przyjmowane obecnie wybory są uzasadnione historycznie i niekiedy są wymagane przez inne działy teorii. Na przykład, gdybyśmy uznali za cząstki proton i antyneutron, to te cząstki nie mogłyby być elementami tego samego dubletu izospinowego.

Cząstki identyczne ze swoimi antycząstkami nazywa się istotnie, lub całkowicie, neutralnymi. Ponieważ sprzężenie ładunkowe zmienia znak ładunku elektrycznego cząstki, warunkiem koniecznym istotnej neutralności jest, żeby cząstka miała zerowy ładunek elektryczny, ale nie jest to warunek wystarczający. Na przykład, neutron ma ładunek elektryczny zero, ale nie jest identyczny ze swoją antycząstką – neutron i antyneutron mogą anihilować ze sobą, a dwa neutrony nie mogą. Wśród bozonów, przykładami cząstek istotnie neutralnych są foton i mezon π^0 . Nie wiadomo, czy istnieją istotnie neutralne fermiony. Ze znanych fermionów jedynymi kandydatami są neutrina. Wróćmy do tego zagadnienia w paragrafie 5.11.

Badanie układów wielocząstkowych wychodzi poza ramy relatywistycznej mechaniki kwantowej, ale dla cząstek nie oddziałujących między sobą działa prosta reguła, że sprzężenie ładunkowe wielocząstkowej funkcji falowej można otrzymać dokonując sprzężenia ładunkowego funkcji falowych wszystkich cząstek wchodzących w skład układu.

Ćwiczenie Pokazać, że układ złożony z n mezonów π^+ , n mezonów π^- i n_0 mezonów π^0 może, ale nie musi, być stanem własnym operatora sprzężenia ładunkowego.

•••••

Cząstki całkowicie neutralne mogą być produkowane i anihilowane pojedynczo, więc ich parzystość ładunkową można wyznaczyć doświadczalnie, jeśli taki proces produkcji idzie przez oddziaływania silne lub elektromagnetyczne.

Jeżeli cząstka jest różna od swojej antycząstki, to czynnik fazowy ϵ_C jest kwestią konwencji. Można go położyć równym $+1$, ale częściej przyjmuje się, że jeśli cząstka należy do multipletu, który zawiera cząstkę całkowicie neutralną, to faza jest taka jak dla tej cząstki. W przypadku mezonów π obie konwencje są równoważne.

Odwrócenie czasu Operator odwrócenia kierunku czasu oznaczamy \hat{T} . Z definicji

$$\hat{T}x = \{-t, \mathbf{x}\}. \quad (3.41)$$

Przyjmujemy, że dla bozonów można położyć $\hat{T}^2 = 1$ Dla cząstek o spinie zero mamy więc, uwzględniając antyliniowość tego operatora,

$$\hat{T}\phi(x) = \epsilon_T \phi^*(\hat{T}x). \quad (3.42)$$

Transformacja jest antyliniowa, więc nie ma ograniczeń na czynnik fazowy ϵ_T . Wobec

$$\hat{T}e^{-iEt} = (e^{-iE(-t)})^* = e^{-iEt}, \quad (3.43)$$

odwrócenie czasu nie zmienia cząstki w antycząstkę. Tak powinno być, więc mamy tu jeszcze jeden argument za antyliniowością operatora odwrócenia czasu.

Zmiana znaku t , podobnie jak sprzężenie zespolone, nie zmienia operatora $(\partial_\mu \partial^\mu + m^2)$, więc odwrócenie czasu jest symetrią równania Kleina-Gordona dla cząstki swobodnej. Podobnie jak inwersja i sprzężenie ładunkowe, ta symetria zachowuje się po włączeniu oddziaływań elektromagnetycznych i silnych, natomiast jest łamana przez oddziaływania słabe. To łamanie jest jednak znacznie mniej widoczne niż dla poprzednich dwu symetrii. Na przykład dla neutrina, odwrócenie czasu zmienia znak zarówno spinu jak i pędu cząstki, a więc neutrino lewoskrętne przechodzi w neutrino prawoskrętne, co nie wskazuje na to, że symetria jest łamana.

Symetrie CPT i CP Jeśli jakieś operatory są dobrymi symetriami, to jest nią także ich iloczyn. Może się jednak zdarzyć, że iloczyn jest dobrą symetrią, choć poszczególne czynniki nie są symetriami, albo że iloczyn jest znacznie bliższy dobrej symetrii niż czynniki. Przykładem pierwszej sytuacji jest operator $\hat{C}\hat{P}\hat{T}$. Tak zwane **twierdzenie CPT** orzeka, że ta symetria jest dobra w każdej teorii pól kwantowych, która spełnia następujące trzy warunki⁴:

- Teoria jest kowariantna względem transformacji $L_{\uparrow+}$.
- Gęstości oddziaływania są lokalne, to znaczy zbudowane z pól kwantowych i ich pochodnych skończonych rzędów w danym punkcie.
- Gęstości są hermitowskie, zsymetryzowane lub zantysymetryzowane według zwykłego związku między spinem i statystyką.

Te założenia są tak ogólne, że wnioski z twierdzenia *CPT*, jak równość mas czy czasów życia cząstek i odpowiadających im antycząstek, uchodzą za szczególnie pewne. Oba powyższe wnioski łatwo wynikają z symetrii względem sprzężenia ładunkowego, ale ta symetria jest tylko przybliżona, więc i równości byłyby przybliżone. Na mocy wyprowadzenia z symetrii *CPT*, równości powinny być dokładne. Może w bardzo silnych polach grawitacyjnych uda się kiedyś zaobserwowanie łamania symetrii *CPT*, bo kwantowa teoria grawitacji, której jeszcze nie ma, prawie na pewno nie będzie spełniała założeń twierdzenia.

Przykładem symetrii przybliżonej, ale znacznie lepszej niż składające się na nią symetrie, jest **symetria CP** odpowiadająca operatorowi $\hat{C}\hat{P}$. Zauważmy, że ta symetria zamienia lewoskrętne neutrino w prawoskrętne antyneutrino, co nie przeczy doświadczeniu. Już od czasu doświadczalnego odkrycia, w roku 1964, łamania symetrii *CP* w słabych oddziaływaniach uchodziło za pewne, na mocy twierdzenia *CPT*, że łamana jest także symetria względem odwrócenia czasu, chociaż bezpośredni, doświadczalny dowód łamania tej symetrii został uzyskany dopiero w roku 2012. Badanie łamania symetrii *CP* jest obecnie ważną dziedziną w fizyce cząstek.

⁴Dyskusję symetrii CPT i dowód twierdzenia CPT można znaleźć, na przykład, w monografiach [8], [11]. Sformułowanie założeń przyjmuję podobne do podanego w [11].

Rozdział 4

Spinory w szczególnej teorii względności

4.1 Grupa $SL(2, C)$, spinory z górnymi wskaźnikami

W fizyce nierelatywistycznej funkcja falowa cząstki o spinie połówkowym jest spinorem (względem grupy obrotów). Podobnie w fizyce relatywistycznej do opisu cząstki o spinie połówkowym potrzebne są spinory (względem grupy $L_{\uparrow+}$)¹.

Spinory względem obrotów w trójwymiarowej przestrzeni euklidesowej są wektorami należącymi do dwuwymiarowej przestrzeni, niezmienniczej i nieredukowalnej względem grupy obrotów. Wygodnie jest zastąpić grupę obrotów grupą dwuwymiarowych macierzy unitarnych o wyznaczniku równym jeden, czyli grupą $SU(2)$. Dla obrotów o kąty mniejsze niż 2π ta grupa działa dokładnie jak grupa obrotów, ale ma wszystkie reprezentacje jednoznaczne, podczas gdy grupa obrotów, dla spinów połówkowych, ma reprezentacje dwuznaczne (obrót o kąt 2π zmienia znak spinora). Podobnie grupę $L_{\uparrow+}$ można zastąpić grupą $SL(2, C)$, czyli grupą dwuwymiarowych macierzy o elementach zespolonych i wyznaczniku równym jeden. Symbol L oznacza, że chodzi o transformacje liniowe, czyli przedstawialne jako mnożenie przez macierz. Dalej ważną będzie uwaga, że grupa $SU(2)$ jest podgrupą grupy $SL(2, C)$, podobnie jak grupa obrotów jest podgrupą grupy Lorentza. Dla transformacji bliskich identyczności, które będziemy głównie dyskutować, można zastępować grupę $SL(2, C)$ przez grupę $L_{\uparrow+}$.

Oznaczając spinory przez ξ i macierze z grupy $SL(2, C)$ przez A mamy wzór na transformację składowych spinora

trabaz

$$\xi'^A = A^A_B \xi^B. \quad (4.1)$$

Przyjmujemy, że wskaźniki oznaczone dużymi literami z początku alfabetu łacińskiego przybierają wartości $\{1, 2\}$ i zachowujemy konwencję sumowania po powtórzonych wskaźnikach, z których jeden jest dolny, a drugi górny. Wskaźnik przy ξ numeruje składowe spinora.

¹Fizycy niekiedy nazywają spinory względem grupy obrotów spinorami Pauli'ego, a spinory względem grupy $L_{\uparrow+}$ spinorami Weyla.

Żeby przejść do innej bazy w przestrzeni spinorów, mnożymy obie strony równości (4.1) przez jakąkolwiek stałą, dwuwymiarową macierz U , posiadającą macierz odwrotną U^{-1} . Wynik może być zapisany w postaci macierzowej jako

$$U\xi' = UAU^{-1}U\xi. \quad (4.2)$$

Porównując ze wzorem (4.1) widać, że macierz A i spinor ξ zostały zastąpione odpowiednio przez

$$\tilde{A} = UAU^{-1} \quad \text{i} \quad \tilde{\xi} = U\xi. \quad (4.3)$$

Reprezentacje A i \tilde{A} powiązane tym wzorem nazywamy równoważnymi, bo działają w tej samej przestrzeni. Wiadomo, że wszystkie dwuwymiarowe reprezentacje grupy obrotów są równoważne. Jak zobaczymy, dla grupy $SL(2, C)$ sytuacja jest bardziej skomplikowana, bo pojawiają się dwie nierównoważne dwuwymiarowe reprezentacje grupy, i co za tym idzie dwie różne dwuwymiarowe przestrzenie nieredukowalne, i dwa rodzaje spinorów.

Wybierając bazę jak dla grupy obrotów ($j = \frac{1}{2}$, $m = \pm\frac{1}{2}$), przy pomocy współczynników Clebscha-Gordana możemy zbudować wyrażenie dwuliniowe, niezmiennicze względem obrotów: $\xi^1\eta^2 - \xi^2\eta^1$. Przy transformacjach z grupy $L_{\uparrow+}$ to wyrażenie transformuje się według wzoru

$$(\xi^1\eta^2 - \xi^2\eta^1)' = (A^1_1\xi^1 + A^1_2\xi^2)(A^2_1\eta^1 + A^2_2\eta^2) - (A^2_1\xi^1 + A^2_2\xi^2)(A^1_1\eta^1 + A^1_2\eta^2). \quad (4.4)$$

Po wykonaniu mnożeń i uporządkowaniu można ten wzór sprowadzić do postaci

$$(\xi^1\eta^2 - \xi^2\eta^1)' = (\xi^1\eta^2 - \xi^2\eta^1) \det A = \xi^1\eta^2 - \xi^2\eta^1, \quad (4.5)$$

gdzie $\det A$ jest wyznacznikiem z macierzy A . Skorzystaliśmy tu z warunku:

$$\det A = 1. \quad (4.6)$$

Zauważmy, że bez tego warunku nie mielibyśmy w teorii skalarów, bo tylko wielkości skalarne względem grupy obrotów mają szansę być skalarne względem grupy $L_{\uparrow+}$.

Dla transformacji bardzo bliskich identyczności możemy napisać

$$A = 1 + c + (\mathbf{a} + i\mathbf{b})\boldsymbol{\sigma}. \quad (4.7)$$

W tym wzorze c jest liczbą zespoloną, wektory \mathbf{a} i \mathbf{b} są rzeczywiste a σ_i , $i = x, y, z$, są macierzami Pauliego. Każdą macierz 2×2 -wymiarową da się zapisać w tej postaci, ale nam wystarczy wzór dla macierzy bliskich jedności, kiedy $|c|$, $|a_i|$ i $|b_i|$ są bardzo małe. Z warunku (4.6) otrzymujemy wtedy

$$c = 0. \quad (4.8)$$

Ćwiczenie Sprawdzić, że jeżeli liczby $|c|, |a_i|, |b_i|$ nie są bardzo małe, to z założenia $\det A = 1$ nie wynika, że $c = 0$.

•••••

²Współczynnikami Clebscha-Gordana sprzęga się zwykle wektory bazowe, a nie ich składowe, ale ta różnica nie ma tu znaczenia (ostatnie ćwiczenie z następnego paragrafu)

Dla $\mathbf{a} = \mathbf{0}$ otrzymujemy wzór na macierz obrotu bardzo bliskiego identity. Wektor \mathbf{a} musi więc parametryzować pchnięcia.

Ćwiczenie Pokazać, że macierze $1 + i\mathbf{b}\boldsymbol{\sigma}$ są unitarne i wywnioskować stąd, że odpowiadają obrotom.

•••••

4.2 Spinory z dolnymi wskaźnikami

Podobnie jak dla czterowektorów, wygodnie jest wprowadzić spinory z dolnymi wskaźnikami zdefiniowane tak, żeby zwężenie dawało skalar, czyli żeby

$$\xi^A \eta_A \equiv \xi^1 \eta_1 + \xi^2 \eta_2 = \xi^1 \eta^2 - \xi^2 \eta^1. \quad (4.9)$$

Odczytujemy stąd

podopu

$$\eta_1 = \eta^2; \quad \eta_2 = -\eta^1. \quad (4.10)$$

Możemy teraz wprowadzić macierz ϵ obniżającą wskaźniki

indedo

$$\eta_A = \epsilon_{AB} \eta^B. \quad (4.11)$$

Porównując z poprzednim wzorem otrzymujemy

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -\epsilon^T = -\epsilon^{-1} = (\epsilon^T)^{-1} = i\sigma_y. \quad (4.12)$$

Oznaczmy e_1, e_2 wektory bazowe w przestrzeni spinorów. W takim razie ogólny spinor można zapisać w postaci:

$$\xi = \xi^A e_A. \quad (4.13)$$

Przy zmianie układu odniesienia zmieniają się wektory bazowe e_A i składowe spinora ξ^A , ale spinor ξ powinien pozostać bez zmian. Wynika stąd, że para wektorów bazowych transformuje się jak składowe spinora z dolnymi wskaźnikami z tym, że wskaźnik oznacza numer wektora bazowego, a nie jego składową.

W paragrafie 1.10 pokazaliśmy, że jeżeli macierz po zwężeniu wszystkich jej wskaźników z wskaźnikami wektorów daje niezmiennik, to ta macierz jest tensorem niezmienniczym. To samo rozumowanie stosuje się do tensorów spinorowych. W szczególności z faktu, że

$$\epsilon_{AB} \xi^A \eta^B = \xi^A \eta_A \quad (4.14)$$

jest niezmiennikiem wynika, że ϵ_{AB} jest tensorem spinorowym niezmienniczym³. Tak więc zależnie od potrzeby, możemy interpretować ϵ jako stałą macierz lub jako tensor spinorowy z dwoma wskaźnikami.

Wobec antysymetrii tensora ϵ ,

$$\epsilon_{AB} (\xi^A \eta^B + \xi^B \eta^A) = 0 \quad (4.15)$$

i stąd

$$\xi^A \eta_A = -\xi_A \eta^A. \quad (4.16)$$

³Tensory spinorowe zostaną omówione w paragrafie 4.4.

Zamiana poziomów wskaźników w zwężeniu, która dla czterowektorów nie wpływała na wartość wyrażenia, tu powoduje zmianę znaku. Wynika z tego w szczególności, że kwadrat każdego spinora jest równy zero:

$$\xi^A \xi_A = 0. \quad (4.17)$$

indeup

Podnosząc po obu stronach wzoru (4.11) wskaźnik A i zmieniając poziomy zwężonych wskaźników B otrzymujemy wzór na podnoszenie wskaźnika:

$$\eta^A = -\epsilon^{AB} \eta_B = \epsilon^{BA} \eta_B = \eta_B \epsilon^{BA}, \quad (4.18)$$

gdzie

$$\epsilon^{AB} = \epsilon_{AB}, \quad (4.19)$$

zgodnie z ogólną regułą podnoszenia wskaźników.

Ćwiczenie Sprawdzić, że $\epsilon^{AB} = \epsilon_{AB}$

•••••

Przejście od spinorów z górnymi wskaźnikami do spinorów ze wskaźnikami dolnymi stanowi zmianę bazy. Zobaczmy jak zmienia się przy tym macierz A . Mamy

$$\xi'_A = \epsilon_{AB} \xi'^B = \epsilon_{AB} A^B_C \xi^C = \epsilon_{AB} A^B_C \xi_D \epsilon^{DC} \equiv \tilde{A}_A^B \xi_B. \quad (4.20)$$

atilde

Zdefiniowana tym wzorem macierz \tilde{A} może być zapisana w następujących, równoważnych postaciach

$$\tilde{A} = \epsilon A \epsilon^T = \sigma_y A \sigma_y = (A^{-1})^T. \quad (4.21)$$

tradol

Można też zapisać

$$\xi'_A = \xi_B (A^{-1})^B_A. \quad (4.22)$$

Ćwiczenie Wykorzystać ten wzór do dowodu, że wyrażenie $\xi_A \eta^A$ jest niezmiennikiem.

•••••

Jeżeli przyjmiemy

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{to} \quad \tilde{A} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}. \quad (4.23)$$

Ćwiczenie Jak by się zmieniała przedstawiona tu teoria, gdybyśmy zbudowali wyjściowy niezmiennik sprzęgając współczynniki Clebscha-Gordana spinory z dolnymi wskaźnikami, co miałyby tę zaletę, że zwykle wprowadza się współczynniki Clebscha-Gordana do sprzęgania wektorów bazowych, które transformują się jak spinory z dolnymi wskaźnikami.

•••••

4.3 Spinory ze wskaźnikami kropkowanymi

Biorąc sprzężenie zespolone obu stron równości (4.1), otrzymujemy

$$\xi'^{A*} = A^A_{B*} \xi^{B*}. \quad (4.24)$$

Macierze A^* też spełniają warunek (4.6) i stanowią reprezentację grupy $SL(2C)$. Sprzężenia spinorów w tym wzorze można opuścić, bo dla dowolnego spinora ξ zawsze można znaleźć spinor, który po sprzężeniu zespolonym daje ξ . Tak więc każdy spinor może się transformować albo przy pomocy macierzy A , albo przy pomocy macierzy A^* .

Powstaje pytanie, czy reprezentacje A i A^* są równoważne, to znaczy, czy istnieje taka macierz U , że dla każdej macierzy A

warrow

$$A^* = UAU^{-1}. \quad (4.25)$$

Żeby to zbadać, wystarczy rozpatrywać transformacje bliskie idyntityczności

$$A = 1 + (\mathbf{a} + i\mathbf{b})\boldsymbol{\sigma}, \quad (4.26)$$

$$A^* = 1 + (\mathbf{a} - i\mathbf{b})\boldsymbol{\sigma}^*. \quad (4.27)$$

Warunek (4.25) można przepisać w postaci

$$(\mathbf{a} - i\mathbf{b})\boldsymbol{\sigma}^* = (\mathbf{a} + i\mathbf{b})U\boldsymbol{\sigma}U^{-1}. \quad (4.28)$$

Ta równość musiałaby zachodzić dla dowolnych wektorów \mathbf{a} i \mathbf{b} . Wybierając $\mathbf{b} = \mathbf{0}$ i wektor \mathbf{a} mający tylko jedną składową różną od zera, otrzymujemy kolejno składowe równości

zmbaza

$$\boldsymbol{\sigma}^* = U\boldsymbol{\sigma}U^{-1}. \quad (4.29)$$

Podobnie, wybierając $\mathbf{a} = \mathbf{0}$ i \mathbf{b} z jedną tylko składową różną od zera mamy

zmbazb

$$\boldsymbol{\sigma}^* = -U\boldsymbol{\sigma}U^{-1}. \quad (4.30)$$

Te warunki są ewidentnie między sobą sprzeczne, więc reprezentacja A^* nie jest równoważna z reprezentacją A . Spinory ξ i ξ^* należą do różnych przestrzeni⁴

Ćwiczenie Pokazać, że już sam warunek (4.29) jest niemożliwy do spełnienia.

•••••

Spinory transformujące się przy pomocy macierzy A^* będziemy odróżniać pisząc kropki nad ich wskaźnikami. Tak więc wzór (4.24) przybierze postać

indkro

$$\xi'^{\dot{A}} = A^{\dot{A}}_{\dot{B}} \xi^{\dot{B}}. \quad (4.31)$$

Ponieważ macierz ϵ jest rzeczywista, podnoszenie i opuszczanie kropkowanych wskaźników odbywa się tak jak dla niekropkowanych. Podobnie przebiega też przejście od macierzy A^* do macierzy \tilde{A}^* . Zauważmy, że ponieważ $x = y^*$ jest

⁴Każdy spinor ξ podobnie jak każdy spinor ξ^* można zapisać w postaci $\begin{pmatrix} a \\ b \end{pmatrix}$, ale transformacje Lorentza są w tych przestrzeniach reprezentowane przez nierównoważne ze sobą macierze, więc przestrzenie uważamy za różne.

równoważne z $y = x^*$, jest kwestią konwencji, które spinory nazwiemy kropkowanymi, a które niekropkowanymi. Niezależne od konwencji jest tylko stwierdzenie, że istnieją dwa rodzaje spinorów. Konwencję, którą dalej będziemy używać, wprowadzimy w paragrafie 4.5. Dalej, zamiast spinory z wskaźnikami kropkowanymi (niekropkowanymi) będziemy czasem pisać krócej: spinory kropkowane (niekropkowane).

Zwracamy uwagę na notację. We wzorze (4.24) znaki sprzężenia zespolonego za wskaźnikami oznaczają, że składowe ξ, ξ' i A zostały zastąpione przez ich wartości sprzężone zespolone, przy czym zmieniły się ich własności transformacyjne i nie są już takie jak sugerują wskaźniki. We wzorze (4.31) znak sprzężenia zespolonego stoi przed wskaźnikami A . To znaczy, że macierz jest sprzężona zespolona w stosunku do macierzy A , ale własności transformacyjne są takie jak wynika ze wskaźników. Na przykład

$$\xi^{A*} = \xi^{*\dot{A}}. \quad (4.32)$$

Czasem chcemy zaznaczyć typ spinora ξ , ale nie chcemy pisać wskaźnika, bo mamy na myśli cały spinor, a nie jedną jego składową. W takich przypadkach w miejscu wskaźnika będziemy pisali literę I , odpowiednio z kropką lub bez. Mamy więc cztery typy spinorów: $\xi^I, \xi_I, \xi^{\dot{I}} i \xi_{\dot{I}}$.

Sprawdzimy jeszcze, że dla grupy obrotów spinory niekropkowane i kropkowane dają reprezentacje równoważne. Dla grupy obrotów $\mathbf{a} = \mathbf{0}$, więc warunek (4.29) nie obowiązuje. Warunek (4.30) można przepisać w postaci

$$\sigma^* U = -U \sigma. \quad (4.33)$$

To znaczy, że macierz U komutuje z czysto urojoną macierzą σ_y i antykomutuje z rzeczywistymi macierzami σ_x i σ_z . Taką macierzą jest macierz σ_y pomnożona przez dowolną liczbę. Wybieramy

$$U = i\sigma_y = \epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (4.34)$$

Tak więc, z punktu widzenia własności transformacyjnych spinora pod działaniem operatorów obrotu, kropkowanie wskaźników jest tylko zmianą bazy. Pamiętając że $\epsilon^{AB} = \epsilon_{AB}$, mamy dla obrotów

$$\xi^{\dot{A}} \sim \epsilon_{AB} \xi^B = \xi_A, \quad i \quad \xi_{\dot{A}} \sim -\xi^A, \quad (4.35)$$

gdzie drugi wzór wynika z pierwszego przez obustronną zmianę poziomu wskaźników. Zapis $a \sim b$ oznacza tu, że a i b transformują się tak samo przy obrotach. Ponieważ zmiana znaku spinora nie zmienia jego własności transformacyjnych względem obrotów, znak minus w drugim wzorze można pominąć.

Ćwiczenie Pokazać, że gdyby relacje (4.35) dały się uogólnić na całą grupę $L_{\uparrow+}$, to reprezentacje A i A^* byłyby równoważne.

•••••

W reprezentacji A^* obroty spinorów $\xi^{\dot{I}}$ są reprezentowane przez macierze $(D^{\frac{1}{2}})^*$. Dla spinorów $\xi_{\dot{I}}$, biorąc sprzężenia zespolone we wzorze (4.21), otrzymujemy macierz transformacji

$$\tilde{A}^* = (A^{-1})^\dagger. \quad (4.36)$$

Dla obrotów macierze A są równe unitarnym macierzom $D^{\frac{1}{2}}$, więc macierze $(A^{-1})^\dagger$ też są im równe. Przy obrotach spinory ξ_j transformują się tak samo jak spinory ξ^I , co zgadza się z wzorami (4.35).

Rozważmy dwie cząstki z funkcjami falowymi odpowiednio ξ^I i η_j . Zakładamy, że obie cząstki są w spoczynku, a zatem każda z nich jest dana przez dwie stałe: górną i dolną składową odpowiedniego spinora. Stosunek tych stałych określa, dla każdej cząstki, jej stan spinowy. Jeśli więc obie cząstki są w tym samym stanie spinowym, to ponieważ oba spinory transformują się tak samo przy obrotach, spinory ξ^I i η_j mogą się różnić co najwyżej o stały czynnik. Ponieważ przy pchnięciach spinory ξ^I i η_j transformują się różnie, dla cząstek z prędkościami różnymi od zera, mimo że są tym samym stanie spinowym relacja między tymi spinorami robi się bardziej skomplikowana.

Nasuwa się pytanie, czy nie udałoby się znaleźć dwuwymiarowej reprezentacji nierównoważnej z żadną z dotychczas wprowadzonych. Odpowiedź jest negatywna. Można udowodnić⁵, że każda dwuwymiarowa, nieredukowalna reprezentacja grupy $L_{\uparrow+}$ jest równoważna z reprezentacją A lub A^* .

4.4 Tensory spinorowe

Tensory spinorowe mają się do spinorów tak, jak tensory do wektorów (paragraf 1.10). W szczególności, tensorem spinorowym mającym m wskaźników górnych niekropkowanych i n wskaźników górnych kropkowanych nazywany każdy zbiór 2^{m+n} liczb, które transformuje się pod działaniem transformacji Lorentza według wzoru

spinomn

$$\zeta^{A_1, \dots, A_m, \dot{B}_1, \dots, \dot{B}_n} = A^{A_1}_{C_1} \dots A^{A_m}_{C_m} A^{*\dot{B}_1}_{\dot{D}_1} \dots A^{*\dot{B}_n}_{\dot{D}_n} \zeta^{C_1, \dots, C_m, \dot{D}_1, \dots, \dot{D}_n}, \quad (4.37)$$

co można też zapisać w postaci

$$\zeta^{A_1, \dots, A_m, \dot{B}_1, \dots, \dot{B}_n} \sim \xi_1^{A_1} \dots \xi_m^{A_m} \xi_{m+1}^{\dot{B}_1} \dots \xi_{m+n}^{\dot{B}_n}, \quad (4.38)$$

gdzie znak \sim tym razem oznacza, że obie strony relacji transformują się tak samo pod działaniem transformacji Lorentza. Analogicznie, zamieniając odpowiednie macierze A na \tilde{A} , można zdefiniować tensory spinorowe z pewną liczbą wskaźników dolnych. Zauważmy, że zmiana poziomów wskaźników sprowadza się do zmiany bazy, czyli prowadzi do reprezentacji równoważnej z wyjściową. Dlatego ograniczymy się do dyskusji tensorów spinorowych, które mają wszystkie wskaźniki górne.

Ćwiczenie Pokazać, że kładąc

$$\zeta_{AB} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (4.39)$$

otrzymuje się

$$\zeta^{AB} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \tilde{\zeta}_{AB} = (\zeta^{-1})_{BA} \det \zeta, \quad (4.40)$$

gdzie ζ oznacza macierz ζ_{AB} .

•••••

⁵Dowód można znaleźć w monografii J. Łopuszańskiego [10].

zwezen

Składowe tensora spinorowego ζ , z danymi liczbami wskaźników niekropkowanych i kropkowanych, stanowią bazę przestrzeni wektorowej, niezmienniczej względem grupy transformacji $L_{\uparrow+}$. Oznaczmy tę przestrzeń Z . Powstaje pytanie, kiedy przestrzeń Z jest nieredukowalna? Zauważmy, że każde zwężenie z tensorem niezmienniczym, jak

$$\epsilon_{A_1 A_2} \zeta^{A_1, \dots, A_m, \dot{B}_1, \dots, \dot{B}_n}, \quad (4.41)$$

daje tensor spinorowy, którego składowe stanowią bazę niezmienniczej względem $L_{\uparrow+}$ podprzestrzeni przestrzeni Z . Wynika stąd, że przestrzeń Z jest nieredukowalna tylko, jeśli każde takie zwężenie daje zero. Można udowodnić, że jest to także warunek wystarczający. Zwężenie (4.41) zeruje się, jeśli tensor spinorowy ζ jest symetryczny względem zamiany $A_1 \leftrightarrow A_2$. Uwzględniając wszystkie możliwe zwężenia, wnioskujemy stąd, że **przestrzeń Z jest nieredukowalna, jeśli tensor spinorowy ζ nie zmienia się przy żadnej permutacji wskaźników niekropkowanych ani przy żadnej permutacji wskaźników kropkowanych**. Element tensora spinorowego mającego taką symetrię można jednoznacznie wskazać, przy danych m i n , podając ile wskaźników niekropkowanych jest równych 1 oraz ile wskaźników kropkowanych jest równych 1. Dla wskaźników niekropkowanych takich możliwości jest $m+1$, a dla kropkowanych $n+1$. Stąd wymiar przestrzeni nieredukowalnej wynosi

dimspi

$$d(m, n) = (m+1)(n+1). \quad (4.42)$$

Można pokazać, że **to są wszystkie przestrzenie o skończonym wymiarze, nieredukowalne względem grupy $SL(2C)$** .

Powstaje pytanie, która z tych przestrzeni odpowiada przestrzeni czterowektorów. Szukamy przestrzeni czterowymiarowej i przechodzącej w siebie pod wpływem sprzężenia zespolonego. Pierwszy warunek daje $(m+1)(n+1) = 4$, drugi $m = n$. Przestrzeni czterowektorów odpowiada więc przestrzeń tensorów spinorowych $\zeta^{A\dot{B}}$.

4.5 Związek tensorów spinorowych z czterowektorami

Wyprowadzimy teraz kowariantne relacje między składowymi dowolnego czterowektora p i elementami odpowiadającego mu tensora spinorowego $\zeta^{A\dot{B}}$. Ponieważ, jak pokazaliśmy na końcu poprzedniego paragrafu, przestrzeń tensorów $\zeta^{A\dot{B}}$ jest jedyną, która spełnia te same warunki co przestrzeń czterowektorów, takie relacje muszą istnieć. Najpierw znajdziemy relacje kowariantne względem grupy obrotów, a potem wybierzemy wśród nich te, które są kowariantne względem całej grupy $L_{\uparrow+}$.

Z punktu widzenia grupy obrotów, dwuwskaźnikowy tensor spinorowy odpowiada dwu spinom $\frac{1}{2}$. Z takich dwu spinów można złożyć obiekty T_m^j z $j = 0, 1$. Używając współczynników Clebscha-Gordana, można wyrazić te obiekty przez spinory α_1 i α_2 , które przy obrotach transformują się odpowiednio jak spinory α i β z nierelatywistycznej teorii spinu. Obiekty T_m^j można też interpretować jako operatory tensorowe i wyrazić przez składowe odpowiedniego czterowektora p , który oczywiście na ogół nie jest czteropędem. Rachunek można znaleźć, na przykład, w podręczniku [1]. Otrzymujemy

$$T_1^1 = \frac{1}{\sqrt{2}}(-p_x - ip_y) \sim \alpha_1 \alpha_1 = -\frac{1}{\sqrt{2}}\zeta_{11}, \quad (4.43)$$

$$T_0^1 = p_z \sim \frac{1}{\sqrt{2}}(\alpha_1 \alpha_2 + \alpha_2 \alpha_1) = -\frac{1}{2}(\zeta_{12} + \zeta_{21}), \quad (4.44)$$

$$T_{-1}^1 = \frac{1}{\sqrt{2}}(p_x - ip_y) \sim \alpha_2 \alpha_2 = -\frac{1}{\sqrt{2}}\zeta_{22}, \quad (4.45)$$

$$T_0^0 = \lambda p_0 \sim \frac{1}{\sqrt{2}}(\alpha_1 \alpha_2 - \alpha_2 \alpha_1) = -\frac{1}{2}(\zeta_{12} - \zeta_{21}), \quad (4.46)$$

gdzie wężyk oznacza, że wyrażenia po obu jego stronach transformują się tak samo przy obrotach i gdzie wprowadziliśmy oznaczenie

$$\zeta_{AB} = -\sqrt{2}\alpha_A \alpha_B. \quad (4.47)$$

Wzór dla T_0^0 wymaga komentarza. Czasowa składowa czterowektora nie zmienia się przy obrotach przestrzennych, a zatem transformuje się jak tensor T_0^0 . Natomiast normalizacja i faza tego tensora w stosunku do tensora T_m^1 nie są określone przez teorię grupy obrotów. Dlatego wprowadziliśmy nieokreślony na razie współczynnik λ . Zbieżność używanego tu oznaczenia λ z oznaczeniem stosowanym dla skrętności jest przypadkowa.

Obiekty ζ_{AB} nie spełniają relacji $\zeta_{AB} = \zeta_{BA}$. Można jednak przejść od nich do tak samo transformujących się przy obrotach obiektów $\zeta^{A\dot{B}}$, dla których ta symetria nie obowiązuje. Korzystając ze wzorów na podnoszenie wskaźników spinorowych i ze wzorów (4.35), otrzymujemy:

$$\zeta_{11} = \zeta^2_1 \sim \zeta^{2\dot{1}} \quad (4.48)$$

$$\zeta_{12} = \zeta^2_2 \sim \zeta^{2\dot{2}} \quad (4.49)$$

$$\zeta_{21} = -\zeta^1_1 \sim -\zeta^{1\dot{1}} \quad (4.50)$$

$$\zeta_{22} = -\zeta^1_2 \sim -\zeta^{1\dot{2}}. \quad (4.51)$$

Znak \sim oznacza, że wyrażenia po obu jego stronach transformują się tak samo przy obrotach.

Definiujemy teraz cztery liczby, które na mocy powyższych wzorów, przy każdym wyborze parametru λ , transformują się przy obrotach jak składowe tensora spinorowego:

$$\eta^{2\dot{1}} = p_x + ip_y, \quad (4.52)$$

$$\eta^{2\dot{2}} = -\lambda p_0 - p_z, \quad (4.53)$$

$$-\eta^{1\dot{1}} = \lambda p_0 - p_z, \quad (4.54)$$

$$-\eta^{1\dot{2}} = -p_x + ip_y. \quad (4.55)$$

Biorąc wszystkie możliwe wartości parametru λ , otrzymujemy wszystkie czwórki liczb mające tę własność. Wynika stąd, że dla jakiejś, czy jakichś, wartości parametru λ , $\eta^{I\dot{J}}$ jest tensorem spinorowym względem całej grupy transformacji $L_{\uparrow+}$. Pozostaje znaleźć te wartości λ .

Najpierw zapiszemy $\eta^{A\dot{B}}$ w postaci macierzowej

relusp

$$\eta^{IJ} = \begin{pmatrix} -\lambda p_0 + p_z & p_x - ip_y \\ p_x + ip_y & -\lambda p_0 - p_z \end{pmatrix} = (-\lambda p_0 + \mathbf{p}\boldsymbol{\sigma})^{IJ}. \quad (4.56)$$

tramac

Dowolny tensor spinorowy $\zeta^{A\dot{B}}$, pod działaniem właściwych transformacji Lorentza, transformuje się według wzoru

$$\zeta'^{A\dot{B}} = A^A{}_C A^{*\dot{B}}{}_{\dot{D}} \zeta^{C\dot{D}} = (A\zeta A^\dagger)^{A\dot{B}}. \quad (4.57)$$

Biorąc wyznaczniki z macierzy po obu stronach tej równości, i pamiętając że $\det A = 1$, otrzymujemy

$$\det \zeta'^{IJ} = \det \zeta^{IJ}. \quad (4.58)$$

To znaczy, że wyznacznik z tensora spinorowego ζ^{IJ} jest niezmiennikiem. Warunkiem koniecznym na to, żeby liczby $\eta^{A\dot{B}}$ były składowymi tensora spinorowego jest więc, żeby wyznacznik

$$\det \eta^{IJ} = \lambda^2 (p_0)^2 - \mathbf{p}^2 \quad (4.59)$$

był niezmiennikiem. To zachodzi wtedy i tylko wtedy, gdy $\lambda^2 = 1$, czyli jeśli $\lambda = \pm 1$. Ponieważ są dwie nierównoważne przestrzenie spinorów, wybór znaku λ odpowiada umowie, którą z tych przestrzeni uznamy za przestrzeń spinorów z wskaźnikami niekropkowanymi.

Wybieramy dla spinorów niekropkowanych:

$$\lambda = -1. \quad (4.60)$$

Macierz $\eta^{A\dot{B}}$ dla $\lambda = -1$ jest tensorem spinorowym, który oznaczymy $\zeta^{A\dot{B}}$. Wzór (4.56) daje szukany, wzajemnie jednoznaczny, kowariantny względem transformacji $L_{\uparrow+}$ związek między składowymi tensora spinorowego $\zeta^{A\dot{B}}$ i składowymi czterowektora p . Wiemy jak się transformują czterowektory przy właściwych transformacjach Lorentza. W następnym paragrafie wyprowadzimy stąd wzory transformacyjne dla spinorów.

Ćwiczenie Pokazać, że tylko dla $\lambda^2 = 1$ zwężenie $\eta^{A\dot{B}}\eta_{\dot{B}A}$ jest skalarem.

•••••

Zauważmy jeszcze, że wzór (4.56) dla $\lambda = -1$ można zapisać w postaci

$$\zeta^{A\dot{B}} = (p_\mu \tilde{\sigma}^\mu)^{A\dot{B}}, \quad (4.61)$$

gdzie

$$\tilde{\sigma}^\mu = \{1, -\boldsymbol{\sigma}\}. \quad (4.62)$$

Ponieważ zwężenie tensora ζ^{IJ} z dowolnymi dwoma spinorami ξ_I i η_j daje niezmiennik, z prostego uogólnienia rozumowania z paragrafu 1.10 wynika, że $\tilde{\sigma}^{\mu I\dot{J}}$ jest tensorem niezmienniczym z trzema górnymi wskaźnikami, jednym wektorowym, jednym spinorowym niekropkowanym i jednym spinorowym kropkowanym. Przystawiając wskaźniki w tensorze $\zeta^{A\dot{B}}$, a następnie obniżając je, otrzymujemy

$$\zeta_{\dot{B}A} = \epsilon_{\dot{B}\dot{D}}\epsilon_{AC}\zeta^{\dot{D}C} = \epsilon_{\dot{B}\dot{D}}\zeta^{\dot{D}C}(\epsilon^T)_{CA}. \quad (4.63)$$

Korzystając z tożsamości $\epsilon = i\sigma_y$, możemy ten wzór przepisać w postaci macierzowej jako

$$\zeta_{IJ} = \sigma_y(p_0 + p_x\sigma_x + p_y\sigma_y^T + p_z\sigma_z)\sigma_y = p_\mu\sigma^\mu, \quad (4.64)$$

gdzie

$$\sigma_{\dot{B}A}^\mu = \{1, \boldsymbol{\sigma}\}_{\dot{B}A}, \quad (4.65)$$

jest znów tensorem niezmienniczym.

4.6 Właściwe transformacje Lorentza spinorów

Każdą transformację $L_{\uparrow+}$ można wykonać w trzech etapach: obrót, który ustawia oś z równoległe do kierunku pchnięcia, pchnięcie bezobrotowe wzdłuż osi z i jeszcze jeden obrót, który ustawia osie x, y, z w ich końcowym położeniu. Ponieważ własności transformacyjne spinorów przy obrotach przyjmujemy za znane, wystarczy przedyskutować pchnięcia bezobrotowe wzdłuż osi z . Podobnie jak dla grupy obrotów, wygodnie jest zacząć od transformacji bliskich identyczności, czyli od bezobrotowych pchnięć z bardzo małą prędkością wzdłuż osi z . Dla takich pchnięć, macierz transformacji

$$A = 1 + \mathbf{a}\boldsymbol{\sigma} = A^\dagger. \quad (4.66)$$

Uwzględniliśmy tu fakt, że $\mathbf{b} = \mathbf{0}$, jeśli nie ma obrotu. Naszym zadaniem jest znalezienie macierzy A , lub równoważnie wektora \mathbf{a} , jako funkcji prędkości pchnięcia v .

Tensor spinorowy ζ transformuje się według wzoru (4.57), więc z dokładnością do członów liniowych w \mathbf{a} :

trzeta

$$\zeta' = \zeta + (\mathbf{a}\boldsymbol{\sigma})\zeta + \zeta(\mathbf{a}\boldsymbol{\sigma}). \quad (4.67)$$

Tu i w dalszym ciągu nie wypisujemy już wskaźników spinorowych. Tensor ζ do końca paragrafu oznacza ζ^{IJ} .

Porównamy dwa wyrażenia na tensor $\zeta' - \zeta$. Z powyższego wyrażenia, podstawiając definicję (4.56), ze zmianą oznaczenia η na ζ , mamy

$$\zeta' - \zeta = (\mathbf{a}\boldsymbol{\sigma})(p_0 + \mathbf{p}\boldsymbol{\sigma}) + (p_0 + \mathbf{p}\boldsymbol{\sigma})(\mathbf{a}\boldsymbol{\sigma}) = 2p_0(\mathbf{a}\boldsymbol{\sigma}) + 2\mathbf{a}\mathbf{p}, \quad (4.68)$$

gdzie skorzystaliśmy z reguły antykomutacji $\sigma_j\sigma_k + \sigma_k\sigma_j = 2\delta_{jk}$. Natomiast wprost z definicji (4.56)

$$\zeta' - \zeta = p'_0 - p_0 + (\mathbf{p}' - \mathbf{p})\boldsymbol{\sigma}. \quad (4.69)$$

Przyrównując do siebie te dwa wyrażenia i korzystając ze wzorów na transformację czterowektora przy bezobrotowym pchnięciu z bardzo małą prędkością v równoległe do osi z :

$$p'_0 - p_0 = vp_z, \quad p'_x - p_x = 0, \quad p'_y - p_y = 0, \quad p'_z - p_z = vp_0, \quad (4.70)$$

otrzymujemy:

$$a_x = a_y = 0, \quad vp_z = 2a_z p_z, \quad vp_0 = 2p_0 a_z. \quad (4.71)$$

Dwie ostatnie równości są równoważne i dają $a_z = \frac{1}{2}v$. Mamy więc w końcu, dla pchnięć wzdłuż osi z bardzo małą prędkością v :

$$A = 1 + \frac{v}{2}\sigma_z. \quad (4.72)$$

Ćwiczenie Powtórzyc to wyprowadzenie, używając zamiast wszystkich składowych czterowektora p ich kombinacje liniowe $p_0 \pm p_z$. Uzyskać wzór na macierz A z prędkością zastąpioną przez pospieszność. Uogólnić wyprowadzenie na pchnięcia wzdłuż osi z ze skończoną prędkością.

•••••

Znajdziemy teraz macierz A dla pchnięć ze skończoną prędkością wzdłuż osi z . W teorii grupy obrotów można było skończony obrót o kąt ϕ przedstawić jako wynik dużej liczby, N , obrotów o bardzo małe kąty $\frac{\phi}{N}$. Tu nie możemy bezpośrednio przenieść tej metody, bo relatywistyczne dodawanie prędkości nie jest liniowe. Można jednak bardzo małą prędkość v zastąpić równą jej z dostateczną dokładnością bardzo małą pospiesznością $y = \text{artanh } v$ (paragraf 1.4). Pośpieszności y są addytywne, więc możemy napisać dla skończonej wartości y i bardzo dużego N :

$$A(y) = A^N\left(\frac{y}{N}\right) = \left(1 + \frac{y\sigma_z}{2N}\right)^N = e^{\frac{y}{2}\sigma_z}. \quad (4.73)$$

Wektory $u_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ i $u_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ są wektorami własnymi macierzy A do wartości własnych odpowiednio $e^{y/2}$ i $e^{-y/2}$, więc dla pchnięcia z pospiesznością y wzdłuż osi z otrzymujemy:

$$A(y) = A^*(y) = \begin{pmatrix} e^{\frac{y}{2}} & 0 \\ 0 & e^{-\frac{y}{2}} \end{pmatrix}. \quad (4.74)$$

Ćwiczenie Udowodnić to wynikanie.

•••••

Dyskusję zaczniemy od spinorów niekropkowanych z górnymi wskaźnikami. Niech wektory u_{\pm} będą spinorowymi funkcjami falowymi cząstki w spoczynku, która ma rzut spinu na oś z równy odpowiednio $\pm\frac{1}{2}$. Dla dowolnych stałych współczynników c_{\pm} :

$$A(y)(c_+u_+ + c_-u_-) = e^{\frac{y}{2}}c_+u_+ + e^{-\frac{y}{2}}c_-u_-. \quad (4.75)$$

Załóżmy, że układ spoczynkowy jest układem laboratoryjnym. Wtedy cząstka poruszająca się z dużą dodatnią prędkością wzdłuż osi z , ($y \rightarrow \infty$) jest widziana

aodeta

z układu laboratoryjnego jako prawoskrętna, jeśli tylko $c_+ \neq 0$. Ponieważ wybór osi z jest dowolny, wynik przenosi się na wszystkie bardzo szybkie cząstki, których funkcje falowe są spinorami niekropkowanymi⁶.

Ćwiczenie Przedyskutować podobnie przypadek $y \rightarrow -\infty$.

•••••

Korzystając ze wzoru (4.36), otrzymujemy macierze reprezentujące pchnięcia dla spinorów kropkowanych z dolnymi współczynnikami

aodeta

$$\tilde{A}^*(y) = \begin{pmatrix} e^{-\frac{y}{2}} & 0 \\ 0 & e^{\frac{y}{2}} \end{pmatrix}. \quad (4.76)$$

Rozumując jak dla spinorów niekropkowanych, wnioskujemy że obserwator spoczywający w układzie laboratoryjnym widzi bardzo szybką cząstkę, opisywaną przez spinor kropkowany, jako lewoskrętną,

Ćwiczenie Wyjaśnić dlaczego ze wzoru $A(y) = A^*(y)$ nie wynika, że szybka cząstka opisywana przez spinor kropkowany jest widziana jako prawoskrętna.

•••••

4.7 Inwersja przestrzenna

Wpływ operatora inwersji przestrzennej \hat{P} na punktochwile i cząstki skalarne względem właściwych transformacji Lorentza dyskutowaliśmy już w paragrafie 3.3. Zwróciliśmy tam uwagę, że z warunku $\hat{P}^2 x = x$ nie wynika, że reprezentant operatora \hat{P}^2 działający na spinory też jest równoważny z operatorem jednostkowym. Teraz potrzebny nam będzie ten reprezentant. Warunek, że układ współrzędnych nie zmienia się, pozostawia możliwość, że dokonujemy obrotu wokół dowolnej osi o kąt 2π . Przy takim obrocie spinor zmienia znak, a zatem nie można wykluczyć, że

pekwad

$$\hat{P}^2 \xi(x) = -\xi(\hat{P}x). \quad (4.77)$$

Okazuje się, że znak tego reprezentanta \hat{P}^2 wpływa na jedno tylko przewidywanie teorii. Przy wyborze (4.77) jest możliwe, żeby fermion był identyczny ze swoją antycząstką. Przy przeciwnym wyborze jest to niemożliwe. Obecnie uchodzi za bardzo prawdopodobne, że antyneutrino jest identyczne z neutrinem, wrócimy do tej hipotezy omawiając teorię Majorany. Wykluczanie takiej możliwości wydaje się więc nierozsądne, i od jakiegoś czasu przyjmowany jest wzór (4.77). Dobrze jest jednak wiedzieć, że używanie przeciwnego znaku, jak to jest na ogół robione w starszych podręcznikach, też prowadzi zwykle do poprawnych wyników.

Przy dyskusji dla cząstek z różnym od zera spinem istotne jest, że inwersja przestrzenna zmienia prędkość cząstki na przeciwną, natomiast nie zmienia orientacji spinu cząstki, w szczególności zachowuje rzut spinu na oś z . Tak więc mamy

spcons

$$\hat{P}\hat{L}(\mathbf{v}) = \hat{L}(-\mathbf{v})\hat{P} \quad \hat{P}S_z\hat{P}^{-1} = S_z. \quad (4.78)$$

⁶Skretność cząstki jest wielkością mierzalną, a więc nie zależy od reprezentacji. Tu wygodnie jest wybrać spinory niekropkowane z górnymi wskaźnikami i kropkowane z dolnymi.

twopos

Zmienia się przy tym skrętność cząstki. W teorii relatywistycznej mamy dwa rodzaje spinorów, i w związku z tym dwie możliwości zachowania z -owej składowej spinu:

$$\hat{P}\xi^A(x) = \epsilon_P \xi^A(\hat{P}x) \quad \text{lub} \quad \hat{P}\xi^A(x) = i\eta_A(\hat{P}x). \quad (4.79)$$

Ponieważ orientacja osi z jest dowolna, współczynnik ϵ_P nie może zależeć od A .

loropa

Żeby dokonać wyboru między możliwościami (4.79) wykorzystamy następujące twierdzenie. Jeśli dla jakiegoś operatora \hat{A} spełniona jest relacja

$$\hat{A}\hat{L}(v) = \hat{L}(-v)\hat{A}, \quad (4.80)$$

gdzie prędkość pchnięcia jest skierowana wzdłuż osi z , to musi zachodzić

$$\hat{A} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ c \end{pmatrix}, \quad (4.81)$$

gdzie c jest jakąś liczbą. Kropkowanie spinorów nie ma tu znaczenia, natomiast ich wskaźniki muszą być na tej samej wysokości. Dla dowodu zauważmy, że jeśli

$$\hat{A} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad (4.82)$$

to

$$\hat{A}\hat{L}(v) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e^{\frac{v}{2}} \hat{A} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e^{\frac{v}{2}} \begin{pmatrix} a \\ b \end{pmatrix} \quad (4.83)$$

oraz

$$\hat{L}(-v)\hat{A} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \hat{L}(-v) \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} e^{-\frac{v}{2}} a \\ e^{\frac{v}{2}} b \end{pmatrix}. \quad (4.84)$$

Te dwa wzory są niesprzeczne ze sobą, tylko jeśli $a = 0$, co było do udowodnienia. Operator \hat{P} spełnia założenie (4.80), więc musimy odrzucić pierwszą z możliwości (4.79).

Ćwiczenie Pokazać, że w pierwszym ze wzorów (4.79) nie można po prawej stronie zastąpić składowych spinora ξ składowymi jakiegoś innego spinora ζ .

Wskazówka: Zacząć od stanów spinowych będących wektorami własnymi macierzy S_z .

•••••

spainv

Przyjmując drugą z możliwości (4.79) mamy następujące wzory:

$$\begin{aligned} \hat{P}\xi^A(x) &= i\eta_A(\hat{P}x), & \hat{P}\eta_A(x) &= i\xi^A(\hat{P}x), \\ \hat{P}\xi_A(x) &= -i\eta^{\hat{A}}(\hat{P}x), & \hat{P}\eta^{\hat{A}}(x) &= -i\xi_A(\hat{P}x). \end{aligned} \quad (4.85)$$

Analogiczne czwórki wzorów spotkamy w następnych dwu paragrafach, więc warto przeprowadzić ogólną dyskusję współczynników fazowych. Fazy we wzorach trzecim i czwartym otrzymujemy zmieniając poziomy wskaźników w pierwszych dwu wzorach. W pierwszym wzorze faza może być dowolna, ale jest ukryta w definicji spinora η_j . Odpowiada to sytuacji omówionej przy dyskusji równania Kleina-Gordona – w działaniu na różne cząstki operator \hat{P} może dawać różne

czynniki fazowe. Pozostaje do ustalenia faza w drugim wzorze. To wymaga przyjęcia dodatkowego warunku. Przyjmujemy warunek (4.77). Przy tym założeniu, działając na obie strony pierwszej równości (4.85) operatorem $-i\hat{P}$, otrzymujemy drugą równość.

Tensory spinorowe transformują się przy inwersji przestrzennej jak iloczyny odpowiednich spinorów.

4.8 Sprzężenie spinora i operacja CP

Operator sprzężenia oznaczamy \hat{C} . Jego działanie na spinory jest dane wzorami *spicon*

$$\begin{aligned}\hat{C}\xi^A(x) &= \epsilon_C \eta^{*A}(x), & \hat{C}\eta_A(x) &= \epsilon_C \xi_A^*(x), \\ \hat{C}\xi_A(x) &= \epsilon_C \eta_A^*(x), & \hat{C}\eta^A(x) &= \epsilon_C \xi^{*A}(x).\end{aligned}\quad (4.86)$$

Spinor η^I jest taki jak dla inwersji przestrzennej, ale ponieważ stały współczynnik fazowy został już wybrany, dla sprzężenia wprowadzamy nowy współczynnik i oznaczamy go ϵ_C . Jako dodatkowy warunek przyjęliśmy:

$$\hat{C}^2 = 1. \quad (4.87)$$

Wtedy wybór fazy w czwartym wzorze, a zatem i w drugim, jest poprawny, bo daje

$$\hat{C}^2 \xi^A(x) = \hat{C} \left(\epsilon_C \eta^{*A}(x) \right) = |\epsilon_C|^2 \xi^A(x). \quad (4.88)$$

Dyskutujemy sprzężenie, a nie sprzężenie ładunkowe, bo w teorii spinorów ładunek nie występuje, ale definicje wybierane są tak, żeby przy dyskusji fizycznych równań można było łatwo przejść do sprzężenia ładunkowego.

Zauważmy, że operator \hat{C} działa zarówno na współczynniki przy wektorach bazowych, jak i na same wektory bazowe, ale działa różnie. Na przykład, jeśli wektorem bazowym jest wektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, to

$$\hat{C} \begin{pmatrix} a \\ 0 \end{pmatrix}^I = \hat{C} \left[a \begin{pmatrix} 1 \\ 0 \end{pmatrix}^I \right] = \epsilon_C a^* \begin{pmatrix} 1 \\ 0 \end{pmatrix}^I = \epsilon_C \begin{pmatrix} a^* \\ 0 \end{pmatrix}^I. \quad (4.89)$$

Współczynnik został sprzężony, a wektor bazowy zmienił własności transformacyjne. Przy innym wyborze wektorów bazowych, sprzężenie działa inaczej. Na przykład, wybierając jako wektor bazowy spinor występujący po lewej stronie powyższego wzoru, mielibyśmy

$$\hat{C} \begin{pmatrix} a \\ 0 \end{pmatrix}^I = \epsilon_C \begin{pmatrix} a \\ 0 \end{pmatrix}^I, \quad (4.90)$$

bo w nowej bazie to ten wektor ma postać $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, a więc nie zmienia się przy sprzężeniu zespolonym. Wykorzystując swobodę wyboru fazy wektorów bazowych, można by wyeliminować czynnik fazowy ϵ_C , ale podobna sytuacja pojawi się w następnym paragrafie przy dyskusji odwrócenia czasu. Ponieważ jednym

wyborem fazy nie da się wyeliminować obu czynników fazowych, nie korzystamy z tej możliwości.

Ćwiczenie Pokazać, że zastępując we wzorach (4.86) wektory bazowe przez te same wektory pomnożone przez $\sqrt{\epsilon_C}$ eliminuje się współczynnik ϵ_C .

•••••

Przy pomocy wzorów (4.85) i (4.86) możemy znaleźć działanie tak zwanej transformacji CP

spicpi

$$\begin{aligned}\hat{C}\hat{P}\xi^A(x) &= -i\epsilon_C\xi_A^*(\hat{P}x), & \hat{C}\hat{P}\eta_{\dot{A}}(x) &= -i\epsilon_C\eta^{*\dot{A}}(\hat{P}x), \\ \hat{C}\hat{P}\xi_A(x) &= i\epsilon_C\xi^{*A}(\hat{P}x), & \hat{C}\hat{P}\eta^{\dot{A}}(x) &= i\epsilon_C\eta_{\dot{A}}^*(\hat{P}x).\end{aligned}\quad (4.91)$$

Znak w pierwszym wzorze jest minus, bo czynnik i , powstały przy działaniu operatora \hat{P} na spinor, ulega sprzężeniu zespolonemu pod działaniem operatora \hat{C} .

Ćwiczenie Udowodnić powyższe wzory.

•••••

Zauważmy, że ta transformacja, w przeciwieństwie do sprzężenia, przeprowadza spinor w spinor należący do tej samej przestrzeni.

4.9 Odwrócenie czasu i operacja CPT

Oznaczmy przez \hat{T} operator odwrócenia czasu. Przypominamy, że ten operator jest antyliniowy i antyunitarny. Pod wpływem odwrócenia czasu

$$\hat{T}x = \{-t, \mathbf{x}\}, \quad \hat{T}\hat{L}(v) = \hat{L}(-v)\hat{T}, \quad \hat{T}S_z\hat{T}^{-1} = -S_z. \quad (4.92)$$

Ponieważ i prędkość, i spin zmieniają zwrot, skrętność nie ulega zmianie.

Mamy dwie możliwości:

dwuznt

$$\hat{T}\xi^A(x) = \epsilon_T\xi^*_A(\hat{T}x) \quad \text{lub} \quad \hat{T}\xi^A(x) = \epsilon_T\eta^{*\dot{A}}(\hat{T}x), \quad (4.93)$$

gdzie ϵ_T jest stałym czynnikiem fazowym. Sprzężenie zespolone składowych spinora ξ wynika z antyliniowości operatora \hat{T} . Na mocy twierdzenia zastosowanego do wybrania jednej z możliwości dla inwersji przestrzennej możemy odrzucić drugi tych wzorów.

Ćwiczenie Pokazać, że w pierwszym ze wzorów (4.93) nie można zastąpić spinora ξ po prawej stronie przez jakiś inny spinor, różniący się od niego bardziej niż o stały czynnik. Zwrócić uwagę na rolę sprzężenia zespolonego w dowodzie.

•••••

Przyjmujemy:

tinvar

$$\begin{aligned}\hat{T}\xi^A(x) &= \epsilon_T\xi^*_A(\hat{T}x), & \hat{T}\eta_{\dot{A}}(x) &= -\epsilon_T\eta^{*\dot{A}}(\hat{T}x), \\ \hat{T}\xi_A(x) &= -\epsilon_T\xi^{*A}(\hat{T}x), & \hat{T}\eta^{\dot{A}}(x) &= \epsilon_T\eta_{\dot{A}}^*(\hat{T}x).\end{aligned}\quad (4.94)$$

Jako dodatkowy warunek przyjęliśmy tu, że spinory, które mają wskaźniki na tym samym poziomie, transformują się tak samo przy odwróceniu czasu. Przy tym założeniu transformacja CPT działa tak samo na wszystkie cztery rodzaje spinorów (4.100). Odwrócenie czasu przeprowadza spinor w spinor tego samego typu.

Z powyższych wzorów wynika, że w działaniu na spinory

tkwadr

$$\hat{T}^2 = -1. \quad (4.95)$$

Na przykład,

$$\hat{T}^2 \xi^A(x) = \hat{T} \left(\epsilon_T \xi_A^*(\hat{T}x) \right) = -|\epsilon_T|^2 \xi^A(x). \quad (4.96)$$

Zauważmy, że trzeci ze wzorów (4.94), który jest tu potrzebny, otrzymuje się z pierwszego przez zmianę poziomów wskaźników, czyli bez żadnych dalszych założeń. Wzór (4.95) jest więc niezależnym od konwencji ogólnym twierdzeniem.

Ponieważ tensor spinorowy transformuje się jak iloczyn spinorów, mamy uogólnienie

$$\hat{T}^2 \zeta^{A_1 \dots A_m \dot{A}_{m+1} \dots \dot{A}_n} = (-1)^n \zeta^{A_1 \dots A_m \dot{A}_{m+1} \dots \dot{A}_n}. \quad (4.97)$$

W grupie obrotów, z n spinów połówkowych możemy złożyć tylko spin całkowity, jeśli n jest parzyste i tylko spin połówkowy (liczba całkowita $+\frac{1}{2}$) jeśli n jest nieparzyste. Wzór (4.95) jest więc poprawny w działaniu na funkcję falową dowolnego układu o połówkowym kręcie całkowitym.

Dla układów o spinie połówkowym, jeśli hamiltonian nie zmienia się przy odwróceniu czasu, wzór (4.95) prowadzi do ciekawego wniosku znanego jako degeneracja Kramersa. Niech funkcja falowa $\psi(x)$ spełnia niezależne od czasu równanie Schrödingera,

$$\hat{H}\psi(x) = E\psi(x), \quad (4.98)$$

Wobec założonego $\hat{T}\hat{H} = \hat{H}\hat{T}$, funkcja $\hat{T}\psi(x)$ też jest rozwiązaniem tego równania. Załóżmy teraz, że poziom energetyczny E jest niezdegenerowany. W takim razie funkcja $\hat{T}\psi(x)$ może się różnić od funkcji $\psi(x)$ co najwyżej o jakiś czynnik fazowy ϵ . Mamy więc

$$\hat{T}^2 \psi(x) = \hat{T} \epsilon \psi(\hat{T}x) = |\epsilon|^2 \psi(x), \quad (4.99)$$

co jest sprzeczne ze wzorem (4.95), jeśli stan opisywany przez funkcję $\psi(x)$ ma połówkowy kręt całkowity. Ta sprzeczność dowodzi, że wszystkie poziomy energetyczne każdego układu o spinie połówkowym są zdegenerowane.

Ćwiczenie Przedyskutować zastosowanie twierdzenia Kramersa do stanów podstawowych atomów wodoru i helu. W przypadku atomu wodoru rozpatrzyć opis bez uwzględnienia spinu protonu i opis z uwzględnieniem tego spinu.

• • • • •

Ćwiczenie Pokazać, że na mocy twierdzenia Kramersa elektron nie może mieć różnego od zera elektrycznego momentu dipolowego (pomijamy tu słabe oddziaływania), natomiast może mieć niezerowy magnetyczny moment dipolowy.

• • • • •

Zauważmy jeszcze, że ponieważ odwrócenie czasu zmienia na przeciwny zarówno spin jak i pęd cząstki, skrętność nie ulega zmianie.

trancp

Wzory (4.85), (4.86) i (4.94) dają transformację CPT:

$$\begin{aligned}\hat{C}\hat{P}\hat{T}\xi^A(x) &= i\epsilon_C\epsilon_T^*\xi^A(-x), & \hat{C}\hat{P}\hat{T}\eta_{\dot{A}}(x) &= -i\epsilon_C\epsilon_T^*\eta_{\dot{A}}(-x), \\ \hat{C}\hat{P}\hat{T}\xi_A(x) &= i\epsilon_C\epsilon_T^*\xi_A(-x), & \hat{C}\hat{P}\hat{T}\eta^{\dot{A}}(x) &= -i\epsilon_C\epsilon_T^*\eta^{\dot{A}}(-x),\end{aligned}\quad (4.100)$$

gdzie wykorzystaliśmy związek $\hat{P}\hat{T}x = -x$. Przy odpowiedniej konwencji faz można otrzymać $\epsilon_C\epsilon_T^* = 1$, i taką konwencję dalej przyjmujemy.

Ćwiczenie Pokazać, że iloczyn $\epsilon_C\epsilon_T^*$ nie zależy od wyboru faz wektorów bazowych.

•••••

4.10 Bispinory

Przestrzeń spinorów niekropkowanych i przestrzeń spinorów kropkowanych są niezmiennicze względem operacji z grupy $L_{\uparrow+}$, ale przy inwersji przestrzennej przechodzą w siebie. Przestrzeń niezmiennicza względem rozszerzonej grupy Lorentza, obejmującej operatory $\hat{L}_{\uparrow+}$ i \hat{P} , musi więc zawierać zarówno spinory ξ^I jak i η_i , a zatem jest czterowymiarowa. Wektorami w tej przestrzeni są tak zwane bispinory

$$\psi(x) = \begin{pmatrix} \xi^I(x) \\ \eta_i(x) \end{pmatrix}. \quad (4.101)$$

Zauważmy, że jak widać ze wzorów (4.35) spinory połączone w bispinor transformują się tak samo pod działaniem grupy obrotów. Oddzielenie spinora z wskaźnikami niekropkowanymi od spinora z wskaźnikami kropkowanymi występuje w reprezentacji, której tu używamy i która jest znana jako **reprezentacja spinorowa**. Po przejściu do innej reprezentacji, bispinor $\psi(x)$ przechodzi w $\psi'(x) = U\psi(x)$, gdzie U może być jakąkolwiek stałą macierzą, 4×4 wymiarową, mającą odwrotność. W bispinorze $\psi'(x)$ na ogół każdy element jest kombinacją liniową elementów spinora $\xi^I(x)$ i elementów spinora $\eta_i(x)$. Wtedy nie da się przypisać kropek poszczególnym elementom bispinora i po prostu się je pomija. Reprezentacja spinorowa jest bardzo wygodna przy dyskusji symetrii i cząstek o wysokiej energii, ale przy niskich energiach wygodniej jest używać innych reprezentacji.

lorbis

Przy właściwych transformacjach Lorentza bispinor w reprezentacji spinorowej transformuje się tak, jak tego wymagają składające się na niego spinory

$$\hat{L}\psi(x) = S\psi(\hat{L}^{-1}x), \quad S = \begin{pmatrix} A & 0 \\ 0 & A^{-1\dagger} \end{pmatrix}. \quad (4.102)$$

Widać, że względem transformacji z grupy $L_{\uparrow+}$ jest to reprezentacja redukowalna. Jest natomiast nieredukowalna względem rozszerzonej grupy Lorentza.

gamma0

Definiujemy 4×4 wymiarową macierz γ^0 , która w reprezentacji spinorowej ma postać

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (4.103)$$

Wprowadziliśmy tu notację, z której dalej będziemy często korzystać. Macierz 4×4 wymiarowa jest przedstawiona jako macierz 2×2 wymiarowa, której elementami są bloki (macierze) 2×2 wymiarowe. Dla macierzy γ^0 bloki diagonalne są macierzami zerowymi, a pozadiagonalne macierzami jednostkowymi. W dalszych rachunkach przyda nam się tożsamość

g0mag0

$$\gamma^0 \begin{pmatrix} a & b \\ c & d \end{pmatrix} \gamma^0 = \begin{pmatrix} d & c \\ b & a \end{pmatrix}, \quad (4.104)$$

gdzie a, b, c, d są dowolnymi macierzami o wymiarze 2×2 .

Ćwiczenie Mnożąc powyższą równość stronami, lewostronnie przez γ^0 i uwzględniając, że $(\gamma^0)^2 = 1$ pokazać, że macierz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ komutuje z macierzą γ^0 jeśli $a = d$ i $b = c$ i antykomutuje, jeśli $a = -d$ i $b = -c$.

•••••

Zauważmy też, że

$$\gamma^0 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a \end{pmatrix}. \quad (4.105)$$

Pomnożenie bispinora przez macierz γ^0 po prostu zamienia miejscami jego górny i dolny spinor. Jako przykład zastosowania wzoru (4.104) otrzymujemy:

sdagsm

$$\gamma^0 S^\dagger \gamma^0 = S^{-1}. \quad (4.106)$$

Działanie operatorów $\hat{O} = \hat{P}, \hat{C}, \hat{T}$ na bispinor powoduje zastąpienia spinorów składowych ξ^I, η_j przez przetransformowane spinory $\hat{O}\xi^I, \hat{O}\eta_j$. Korzystając ze wzorów na transformacje spinorów i ze wzorów na podnoszenie i opuszczanie wskaźników trzymujemy:

$$\psi^P(x) = U_P \psi(\hat{P}x), \quad \text{gdzie} \quad U_P = i\gamma_0, \quad (4.107)$$

$$\psi^C(x) = \epsilon_C U_C \psi^*(x), \quad \text{gdzie} \quad U_C = \begin{pmatrix} 0 & -\epsilon \\ \epsilon & 0 \end{pmatrix}, \quad (4.108)$$

$$\psi^T(x) = \epsilon_T U_T \psi^*(\hat{T}x), \quad \text{gdzie} \quad U_T = \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix}. \quad (4.109)$$

W tych wzorach

$$\psi^*(x) = \begin{pmatrix} \xi^{*I}(x) \\ \eta_j^*(x) \end{pmatrix} \quad (4.110)$$

i wprowadziliśmy oznaczenie

$$\hat{O}\psi(x) = \psi^O(x). \quad (4.111)$$

Podobnie otrzymujemy

$$\psi^{CP}(x) = \epsilon_C U_{CP} \psi^*(x), \quad \text{gdzie} \quad U_{CP} = i \begin{pmatrix} \epsilon & 0 \\ 0 & -\epsilon \end{pmatrix}, \quad (4.112)$$

$$\psi^{CPT}(x) = \epsilon_C \epsilon_T^* U_{CPT} \psi(-x), \quad \text{gdzie} \quad U_{CPT} = i\gamma^5, \quad (4.113)$$

Wprowadziliśmy tu macierz

gamma5

$$\gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (4.114)$$

Powyższe wzory obowiązują w reprezentacji spinorowej. Powstaje pytanie, jak działają operatory \hat{P} , \hat{C} i \hat{T} w innych reprezentacjach, gdzie bispinory $\psi(x)$ są zastąpione przez bispinory $\tilde{\psi}(x) = U\psi(x)$. W tym przypadku

$$\hat{O}\tilde{\psi}(x) = \epsilon_O U' U_O \psi'(\hat{O}x) = U' U_O U^{-1} \tilde{\psi}'(\hat{O}x), \quad (4.115)$$

gdzie prim oznacza sprzężenie zespolone dla $\hat{O} = \hat{C}, \hat{T}$ i powinien być pominięty dla $\hat{O} = \hat{P}$. Zobaczmy jak operator \hat{O} działa na bispinor $\tilde{\psi}$. Dla $O = C, T$, bispinor $\tilde{\psi}$ przechodzi w bispinor $\tilde{\psi}^*$, co oznacza sprzężenie zespolone macierzy U i bispinora ψ . Dla $O = P$ ten krok się pomija. Następnie bispinor ψ^* lub ψ jest przetransformowany zgodnie ze zwykłymi regułami, używając standardowych macierzy U_O i współczynników ϵ_O . Mamy więc

$$\tilde{U}_O = U' U_O U^{-1}. \quad (4.116)$$

Tylko dla $\hat{O} = \hat{P}$, kiedy prim się pomija, macierz \tilde{U}_O jest zawsze równa macierzy U_O w nowej reprezentacji. Dla operatorów $\hat{O} = \hat{C}, \hat{T}$ jest tak tylko, jeśli macierz U jest rzeczywista.

4.11 Chiralność

Użyteczna symetria wiąże się z macierzą γ^5 . Ta macierz jest hermitowska, komutuje z macierzą S i antykomutuje z macierzą γ^0 . Jej kwadrat jest równy jeden, więc jej wektory własne spełniają równania

$$\gamma^5 \psi(x) = \chi \psi(x), \quad \chi = \pm 1. \quad (4.117)$$

Wartość własna χ nazywa się chiralnością bispinora spełniającego to równanie. Bispinor jest prawy, jeśli $\chi = +1$ i lewy, jeśli $\chi = -1$ ⁷.

Z równości

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \xi^I(x) \\ \eta_I(x) \end{pmatrix} = \begin{pmatrix} \xi^I(x) \\ -\eta_I(x) \end{pmatrix} \quad (4.118)$$

widać, że dla stanów prawych $\eta_I(x) \equiv 0$, a więc bispinor transformuje się jak niekropkowany spinor $\xi^I(x)$, a dla stanów lewych $\xi^I(x) \equiv 0$, więc bispinor transformuje się jak kropkowany spinor $\eta_I(x)$. Transformacje \hat{P} i \hat{C} zmieniają kropkowanie spinorów, a więc pod ich działaniem zmienia się na przeciwną chiralność bispinora. Transformacje $\hat{L}_{\uparrow+}$, \hat{T} , $\hat{C}\hat{P}$ i $\hat{C}\hat{P}\hat{T}$ nie zmieniają kropkowania spinorów, a więc zachowują chiralność bispinora.

Każdy bispinor można przedstawić jako kombinację liniową bispinora prawego i bispinora lewego. Pod działaniem operatora

$$P_+ = \frac{1}{2}(1 + \gamma^5), \quad (4.119)$$

⁷Często spotykanymi określeniami bispinorów o określonej chiralności są prawoskrętny lub lewoskrętny, ale nie będziemy tych nazw używać, bo mylą się z określeniami skrętności spinorów.

składowa prawa nie ulega zmianie, a składowa lewa zeruje się. Podobnie, pod wpływem operatora

$$P_- = \frac{1}{2}(1 - \gamma^5), \quad (4.120)$$

składowa lewa nie ulega zmianie, a składowa prawa zeruje się. Operatory P_{\pm} mogą służyć do wydzielenia z dowolnego bispinora odpowiednio jego części prawej lub lewej. Z tego względu są zaliczane do operatorów rzutowych.

4.12 Kowarianty dwuliniowe

W zastosowaniach często występują iloczyny składowych bispinorów

phipsi

$$\phi_i^*(x)\psi_j(x), \quad i = 1, 2, 3, 4, \quad j = 1, 2, 3, 4. \quad (4.121)$$

Zaniedbamy na razie zależność od x . Wtedy można te iloczyny uważać za bazę szesnastowymiarowej przestrzeni, niezmienniczej względem rozszerzonej grupy Lorentza. Ponieważ ta przestrzeń jest redukowalna, wygodnie jest w niej wprowadzić inne wektory bazowe. Ze względu na ich proste własności transformacyjne, i liniowość w ψ i ϕ^\dagger , nazywa się je kowariantami dwuliniowymi. Rozważmy najpierw grupę $\hat{L}_{\uparrow+}$, żeby zobaczyć jakich uproszczeń można się spodziewać.

Oznaczamy

$$\phi^* = \begin{pmatrix} \alpha^I \\ \beta_I \end{pmatrix}, \quad \psi = \begin{pmatrix} \xi^I \\ \eta_I \end{pmatrix}. \quad (4.122)$$

Na skutek sprzężenia zespolonego bispinora ϕ , spinor α zyskał, a spinor β stracił, kropkę nad wskaźnikiem. Mnożąc odpowiednio spinory występujące w bispinorze ϕ^* przez spinory występujące w bispinorze ψ , można utworzyć tensory spinorowe o składowych:

$$\alpha^{\dot{A}}\xi^B, \quad \alpha^{\dot{A}}\eta_{\dot{B}}, \quad \beta_A\xi^B, \quad \beta_A\eta_{\dot{B}}. \quad (4.123)$$

Nie zmieniając przestrzeni, w których te tensory są określone, można dowolnie zmieniać poziomy wskaźników. Ze składowych tensora spinorowego $\alpha^{\dot{A}}\xi^B$ można utworzyć czterowektor. Drugi czterowektor można utworzyć ze składowych tensora $\beta^A\eta^{\dot{B}}$. W pozostałych dwu tensorach podnosimy dolne wskaźniki. Części antysymetryczne uzyskanych w ten sposób dwu tensorów są niezmiennikami. Ich części symetryczne są bazami trójwymiarowych przestrzeni nieredukowalnych, ale wygodniej jest połączyć te przestrzenie w przestrzeń redukowalną — sześciowymiarową przestrzeń antysymetrycznych tensorów $F^{\mu\nu}$. Zauważmy, że dwa skalary, osiem składowych dwu czterowektorów i sześć składowych antysymetrycznego tenora $F^{\mu\nu}$ daje łącznie szesnaście wektorów bazowych tak, jak powinno być.

Wektorów bazowych będziemy szukali w postaci

$$\bar{\phi}\Gamma_i\psi, \quad \bar{\phi} = \phi^\dagger\gamma^0, \quad (4.124)$$

gdzie Γ_i są stałymi macierzami 4×4 wymiarowymi.

Potrzebne nam będą następujące własności transformacyjne funkcji $\bar{\phi}$:

$$\hat{L}\bar{\phi} = (S\phi)^\dagger\gamma^0 = \phi^\dagger S^\dagger\gamma^0 = \phi^\dagger\gamma^0\gamma^0 S^\dagger\gamma^0 = \bar{\phi}S^{-1}, \quad (4.125)$$

$$\hat{P}\bar{\phi} = (i\gamma^0\phi)^\dagger\gamma^0 = -i\phi^\dagger\gamma^0\gamma^0 = -i\bar{\phi}\gamma^0. \quad (4.126)$$

sm1gas

Stąd otrzymujemy

$$\hat{L}\bar{\phi}\Gamma_i\psi = \bar{\phi}S^{-1}\Gamma_i S\psi, \quad (4.127)$$

$$\hat{P}\bar{\phi}\Gamma_i\psi = -i\bar{\phi}\gamma^0\Gamma_i i\gamma^0\psi = \bar{\phi}\gamma^0\Gamma_i\gamma^0\psi. \quad (4.128)$$

W ważnym szczególnym przypadku, kiedy

$$\gamma^0\Gamma_i\gamma^0 = P\Gamma_i, \quad P = \pm 1, \quad (4.129)$$

lornaS

mamy

$$\hat{P}\bar{\phi}\Gamma_i\psi = P\bar{\phi}\Gamma_i\psi \quad (4.130)$$

$$\hat{P}\bar{\phi}\gamma^5\Gamma_i\psi = \bar{\phi}\gamma^0\gamma^5\Gamma_i\gamma^0\psi = -\bar{\phi}\gamma^5\gamma^0\Gamma_i\gamma^0\psi = -P\bar{\phi}\gamma^5\Gamma_i\psi, \quad (4.131)$$

a więc pomnożenie macierzy Γ_i przez γ^5 , nawiasem mówiąc wszystko jedno z której strony, zmienia parzystość wektora bazowego na przeciwną. Pokażemy teraz, że własności transformacyjne pod działaniem operatorów $\hat{L}_{\uparrow+}$ nie ulegają przy tym zmianie. Przyjmijmy, że

$$S^{-1}\Gamma_i S = \sum_j a_{ij}\Gamma_j, \quad (4.132)$$

gamind

gdzie a_{ij} są liczbowymi współczynnikami. Jeżeli macierze Γ_i są liniowo niezależne, to znaczy jeśli równość

$$\sum_j c_j\Gamma_j = 0 \quad (4.133)$$

zachodzi tylko kiedy wszystkie współczynniki c_i są równe zero, to takie rozwiązanie jest zawsze możliwe. Wtedy

$$\hat{L}\bar{\phi}\Gamma_i\psi = \bar{\phi}S^{-1}\Gamma_i S\psi = \sum_j a_{ij}\bar{\phi}\Gamma_j\psi \quad (4.134)$$

$$\hat{L}\bar{\phi}\gamma^5\Gamma_i\psi = \bar{\phi}S^{-1}\gamma^5\Gamma_i S\psi = \bar{\phi}\gamma^5 S^{-1}\Gamma_i S\psi = \sum_j a_{ij}\bar{\phi}\gamma^5\Gamma_j\psi. \quad (4.135)$$

W obu tych wzorach własności transformacyjne są takie same.

Dla pierwszego wektora bazowego wybieramy

$$\Gamma_S = 1. \quad (4.136)$$

W tym przypadku

$$\hat{L}\bar{\phi}\psi = \bar{\phi}S^{-1}S\psi = \bar{\phi}\psi, \quad (4.137)$$

$$\hat{P}\bar{\phi}\psi = \bar{\phi}\gamma^0\gamma^0\psi = \bar{\phi}\psi, \quad (4.138)$$

a więc ten wektor jest skalarem względem wszystkich transformacji rozszerzonej grupy Lorentza. Z poprzedniej dyskusji wynika, że kładąc

$$\Gamma_P = \gamma_5 \quad (4.139)$$

otrzymujemy pseudoskalar. Znajdziemy teraz macierze Γ_V^μ dające kowarianty transformujące się jak składowe czterowektora. Oznaczając je

$$\Gamma_V^\mu = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^\mu, \quad (4.140)$$

otrzymujemy

$$\bar{\phi}\Gamma_V^\mu\psi = \beta_A a_B^{\mu A} \xi^B + \beta_A b^{\mu A\dot{B}} \eta_{\dot{B}} + \alpha^{\dot{A}} c_{AB}^\mu \xi^B + \alpha^{\dot{A}} d_{\dot{A}}^{\mu\dot{B}} \eta_{\dot{B}}. \quad (4.141)$$

Wskaźniki spinorowe elementów macierzy a, b, c, d dobraliśmy tak, żeby można je było zwięzić ze wskaźnikami bispinorów i żeby jedynym niezwięzonym wskaźnikiem został wskaźnik czterowektora μ .

Czterowektor buduje się tylko ze spinorów tensorowych, które mają jeden wskaźnik kropkowany a drugi nie, musi więc zachodzić

$$a = d = 0. \quad (4.142)$$

Ponieważ macierze Γ_V^μ mają być stałe, a wyrażenia $\beta_A b^{\mu A\dot{B}} \eta_{\dot{B}}$ i $\alpha^{\dot{A}} c_{AB}^\mu \xi^B$ mają się transformować jak czterowektory, wielkości $b^{\mu A\dot{B}}$ i c_{AB}^μ muszą być tensorami niezmienniczymi. Korzystając z wyników z końca paragrafu 4.5 możemy przyjąć *Gamma A V*

$$b^\mu = e^{i\omega_b} \{1, -\sigma\}, \quad c^\mu = e^{i\omega_c} \{1, \sigma\}, \quad (4.143)$$

gdzie ω_b i ω_c są dowolnymi rzeczywistymi skalarami z przedziału $[0, 2\pi)$. Ten wybór gwarantuje, że pod działaniem operatorów $\hat{L}_{\uparrow+}$ wielkości $\bar{\phi}\Gamma_V^\mu\psi$ transformują się jak składowe czterowektora. Pozostaje do sprawdzenia działanie operatora inwersji przestrzennej.

Ponieważ

$$\gamma^0 \Gamma_V^\mu \gamma^0 = \begin{pmatrix} 0 & c^\mu \\ b^\mu & 0 \end{pmatrix}, \quad (4.144)$$

ze wzorów (4.143) wynika, że musimy przyjąć $\omega_b = \omega_c$, żeby składowa czasowa czterowektora miała parzystość plus, a pozostałe parzystość minus, tak jak powinno być dla czterowektora. Dla wygody przyjmijmy jeszcze konwencję *gammat*

$$\Gamma_V^\mu \equiv \gamma^\mu = \begin{pmatrix} 0 & \tilde{\sigma}^\mu \\ \sigma^\mu & 0 \end{pmatrix}, \quad (4.145)$$

gdzie wprowadziliśmy często używane oznaczenie γ^μ .

Jak wynika z ogólnej dyskusji, liczby $\bar{\phi}\Gamma_A^\mu\psi$, dla

$$\Gamma_A^\mu = \gamma_5 \gamma^\mu, \quad (4.146)$$

transformują się jak składowe czterowektora aksjalnego.

Wprowadzimy macierze:

$$\Gamma_T^{\mu\nu} = \sigma^{\mu\nu} \equiv \frac{i}{2} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu). \quad (4.147)$$

Macierze $\sigma^{\mu\nu}$ są blisko związane z macierzami Pauli'ego:

$$\sigma^{0i} = -i \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}, \quad \sigma^{ij} = \epsilon_{ijk} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}, \quad (4.148)$$

gdzie ϵ_{ijk} jest tensorem całkowicie antysymetrycznym.

Pokażemy teraz, że wektory $\bar{\phi} \Gamma_T^{\mu\nu} \psi$ transformują się przy właściwych transformacjach Lorentza jak składowe antysymetrycznego tensora.

Z faktu, że wielkości $\bar{\phi} \gamma^\mu \psi$ transformują się jak składowe czterowektora i ze wzorów (1.32), (4.127), wynika, że

$$\bar{\phi} S^{-1} \gamma^\mu S \psi = \Lambda^\mu{}_\nu \bar{\phi} \gamma^\nu \psi = \bar{\phi} \Lambda^\mu{}_\nu \gamma^\nu \psi. \quad (4.149)$$

Ponieważ ta równość zachodzi dla każdej pary $\bar{\phi}, \psi$, wynika z niej

$$S^{-1} \gamma^\mu S = \Lambda^\mu{}_\nu \gamma^\nu. \quad (4.150)$$

Pod wpływem właściwych transformacji Lorentza, macierz $\sigma^{\mu\nu}$ zostaje pomnożona lewostronnie przez macierz S^{-1} i z prawostronnie przez macierz S . Wynik można przepisać następująco:

$$\sigma'^{\mu\nu} = \frac{i}{2} (S^{-1} \gamma^\mu S S^{-1} \gamma^\nu S - S^{-1} \gamma^\nu S S^{-1} \gamma^\mu S). \quad (4.151)$$

Korzystając z poprzedniego wzoru widać, że macierze $\sigma^{\mu\nu}$ rzeczywiście transformują się przy właściwych transformacjach Lorentza jak składowe dwuwskaznikowego tensora antysymetrycznego. Podobnie pokazuje się, że działanie inwersji przestrzennej jest takie jak dla tensora.

Ćwiczenie Pokazać, że pomnożenie macierzy $\sigma^{\mu\nu}$ przez macierz γ^5 nie daje nic nowego, i wyjaśnić tę obserwację.

•••••

Podsumowując, znaleźliśmy 16 macierzy:

$$\Gamma_S = 1, \quad \Gamma_P = \gamma_5, \quad \Gamma_V^\mu = \gamma^\mu, \quad \Gamma_A^\mu = \gamma^5 \gamma^\mu, \quad \Gamma_T^{\mu\nu} = \sigma^{\mu\nu}, \quad (4.152)$$

takich, że liczby $\bar{\phi} \Gamma_i \psi$ transformują się względem rozszerzonej grupy Lorentza odpowiednio jak skalar, pseudoskalar, czterowektor, czterowektor aksjalny i dwuwskaznikowy tensor antysymetryczny. Niezależność tych macierzy wynika z faktu, że każda z nich inaczej się transformuje⁸

Ćwiczenie Pokazać, że gdyby zmienić macierz U_P z $i\gamma^0$ na γ^0 , co by odpowiadało $\hat{P}^2 = 1$, to własności transformacyjne kowariantów dwuliniowych nie uległyby zmianie.

•••••

Dla zależnych od punktochwili bispinorów $\phi(x)$ i $\psi(x)$ używa się tych samych kowariantów dwuliniowych i zachowuje się ich określenia: skalar, pseudoskalar

⁸Algebraik zauważyłby, że macierze Γ_i , podobnie jak macierze σ^μ , stanowią macierzową reprezentację algebry Clifforda. Z tego wynika nie tylko ich liniowa niezależność, ale i to że są określone z dokładnością do podobieństwa, $\Gamma'_i = U \Gamma_i U^{-1}$, przez sam swój wymiar.

itd. Oczywiście, pod działaniem jakiegokolwiek operatora symetrii \hat{A} , oprócz przekształceń dyskutowanych w tym paragrafie, trzeba jeszcze zamienić punktowilę x w punktowilę $\hat{A}^{-1}x$.

4.13 Własności macierzy γ^μ

Ze wzoru (4.145) otrzymujemy wzory na macierze γ^μ w reprezentacji spinorowej: *gamspi*

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\gamma} = \begin{pmatrix} 0 & -\boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}. \quad (4.153)$$

Te macierze są zbudowane z tensorów niezmienniczych, i w tym sensie same też są niezmiennicze.

Jak łatwo sprawdzić mnożąc odpowiednio macierze, macierze γ^μ spełniają reguły antykomutacji: *antkom*

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}. \quad (4.154)$$

Ćwiczenie Korzystając z reguł antykomutacji pokazać, że $\gamma_\mu \gamma^\lambda \gamma^\mu = -2\gamma^\lambda$ i $\gamma_\mu \gamma^\lambda \gamma^\nu \gamma^\mu = 4g^{\lambda\nu}$.

•••••

Powyzsze reguły antykomutacji są tak ważne, że często przyjmuje się je jako definicję macierzy γ^μ . Według twierdzenia udowodnionego przez Pauli'ego⁹, jeśli jakieś inne cztery czterowymiarowe macierze γ'^μ spełniają te same reguły antykomutacji, to istnieje taka macierz U , że *zmirep*

$$\gamma'^\mu = U\gamma^\mu U^{-1}, \quad \text{dla} \quad \mu = 0, 1, 2, 3. \quad (4.155)$$

Są to więc macierze równoważne z macierzami (4.153). Wybór reprezentacji jest kwestią wygody.

Ćwiczenie Znaleźć reprezentacje macierzy γ otrzymane z reprezentacji spinorowej używając $U = \gamma^5$ i $U = \gamma^0$.

•••••

Macierz γ^5 wyraża się przez macierze $\gamma^{\mu 10}$

$$\gamma^5 = i\gamma^0 \gamma^x \gamma^y \gamma^z. \quad (4.156)$$

Z tego wzoru widać, że macierz γ^5 antykomutuje ze wszystkimi macierzami γ^μ :

$$\gamma^\mu \gamma^5 + \gamma^5 \gamma^\mu = 0, \quad \mu = 0, 1, 2, 3. \quad (4.157)$$

Wszystkie macierze Γ_i , wprowadzone w poprzednim paragrafie, można wyrazić przez macierze γ^μ . Zwykle jednak, żeby wzory były bardziej przejrzyste, używa się macierzy γ^μ i γ^5 .

W reprezentacji spinorowej, której używamy, przy sprzężeniu hermitowskim

$$\gamma^{\mu\dagger} = \gamma^\mu. \quad (4.158)$$

⁹Przystępny dowód tego twierdzenia można znaleźć w przypisie C w podręczniku Sakuraia [13]

¹⁰Niektórzy autorzy definiują macierz γ^5 z przeciwnym znakiem.

W praktyce używa się prawie wyłącznie reprezentacji, do których można przejść od reprezentacji spinorowej przy pomocy unitarnych macierzy U . Dla każdej takiej reprezentacji, podobnie jak dla reprezentacji spinorowej:

$$\gamma'^{\mu\dagger} \equiv (U\gamma^\mu U^{-1})^\dagger = U\gamma^{\mu\dagger}U^\dagger = U\gamma_\mu U^{-1} = \gamma'_\mu. \quad (4.159)$$

Często używane jest zaproponowane przez Feynmana oznaczenie

$$\not{a} = a_\mu \gamma^\mu, \quad (4.160)$$

gdzie a jest dowolnym czterowektorem. Symbol po lewej stronie tej równości nazywa się *a przekreślone*.

Ćwiczenie Sprawdzić, że $\not{a}\not{b} + \not{b}\not{a} = 2ab$.

•••••

uctgam

Zapiszemy jeszcze wzory

$$U_C = i\gamma^y, \quad U_T = \gamma^x\gamma^z, \quad U_{CP} = \gamma^0\gamma^y. \quad (4.161)$$

Na podstawie wyników z poprzedniego paragrafu, te wzory stosują się w reprezentacji spinorowej i w każdej reprezentacji $\gamma'^\mu = U\gamma^\mu U^{-1}$, dla której macierz przejścia od reprezentacji spinorowej, U , jest rzeczywista.

Rozdział 5

Równania relatywistyczne dla cząstek o spinie $S = \frac{1}{2}$

5.1 Równania Weyla

W relatywistycznych równaniach dla cząstki swobodnej o spinie $S = \frac{1}{2}$, występuje, oprócz funkcji falowej, tylko czterowektor p^μ , lub równoważnie odpowiadający mu tensor spinorowy z dolnymi wskaźnikami

$$p_{\dot{B}A} = \begin{pmatrix} p_0 - p_z & -p_x + ip_y \\ -p_x - ip_y & p_0 + p_z \end{pmatrix} = (p^0 - \mathbf{p}\boldsymbol{\sigma})_{\dot{B}A} \quad (5.1)$$

czy z górnymi

$$p^{A\dot{B}} = \begin{pmatrix} p_0 + p_z & p_x - ip_y \\ p_x + ip_y & p_0 - p_z \end{pmatrix} = (p^0 + \mathbf{p}\boldsymbol{\sigma})^{A\dot{B}}. \quad (5.2)$$

Równania, zaproponowane przez Weyla w roku 1929 i znane jako równania Weyla, otrzymujemy podstawiając, jak zwykle, za składowe czteropędu operatory $\hat{p}_\mu = i\partial_\mu$ i zakładając, że funkcja falowa jest spinorem. Otrzymamy dwa równania, jedno dla spinorów niekropkowanych, a drugie dla kropkowanych. Ponieważ zarówno inwersja przestrzenna, jak i sprzężenie ładunkowe wymieniają wskaźniki kropkowane i niekropkowane ani inwersja przestrzenna, ani sprzężenie ładunkowe nie są symetriami teorii Weyla. Na tej podstawie do około roku 1957 te równania uchodziły za niefizyczne. W roku 1957 okazało się, że w słabych oddziaływaniach ani parzystość, ani symetria ładunkowa nie są zachowywane. Od tego czasu do około roku 1997 uważano, że neutrino, które oddziałuje tylko słabo i, jak wtedy sądzono, ma masę zero, jest dobrze opisywane przez jedno z równań Weyla. W wielu dobrych podręcznikach z tego okresu równania Weyla są dyskutowane jako teoria neutrino. Jak pokazujemy poniżej, z równań Weyla wynika, że neutrino ma masę dokładnie równą zero. Około roku 1997 wykazano doświadczalnie, że z trzech znanych rodzajów neutrin co najmniej dwa mają masy różne od zera. Ścisłe zerowanie się masy cząstki powinno wynikać z jakichś ogólnych zasad. Trudno sobie wyobrazić zasady, które by wymuszały zerowanie się masy tylko jednego z trzech neutrin, więc znów obecnie uważa się, że równania Weyla nie opisują żadnych znanych cząstek. Niemniej, dyskusja

tych równań stanowi bardzo dobry wstęp do dyskusji innych relatywistycznych równań dla cząstek o spinie połówkowym.

weylnk

Zakładając, że funkcja falowa jest spinorem z wskaźnikiem kropkowanym, otrzymujemy kowariantne równanie

$$\hat{p}^{B\dot{A}}\eta_{\dot{A}}(x) = 0. \quad (5.3)$$

Po prawej stronie musi być zero, bo nie dysponujemy żadnym różnym od zera spinorem z wskaźnikiem niekropkowanym, a równanie musi być kowariantne (względem grupy właściwych transformacji Lorentza). Zmieniając poziomy wskaźników nie dostaje się nic nowego. Dla spinora z wskaźnikiem niekropkowanym otrzymujemy inne równanie Weyla

weylkr

$$\hat{p}_{\dot{B}A}\xi^A(x) = 0. \quad (5.4)$$

To są dwa różne równania, opisujące różne cząstki. Same równania są jednak tak podobne do siebie, że opłaca się rozpatrywać je łącznie. Wprowadzając oznaczenia

$$\zeta_+ = \eta_{\dot{I}}, \quad \zeta_- = \xi^I, \quad (5.5)$$

weylmat

możemy oba równania Weyla w postaci macierzowej zapisać jednym wzorem

$$(\hat{p}^0 \pm \hat{\mathbf{p}}\boldsymbol{\sigma})\zeta_{\pm}(x) = 0. \quad (5.6)$$

Z dyskusji w paragrafie 4.11 wynika, że cząstki opisywane przez spinory ζ_{\pm} mają odpowiednio chiralności $\chi = \mp 1$.

kgweyl

Mnożąc równania Weyla stronami, lewostronnie przez $(\hat{p}^0 \mp \hat{\mathbf{p}}\boldsymbol{\sigma})$ i wykorzystując tożsamość $(\hat{\mathbf{p}}\boldsymbol{\sigma})^2 = \hat{\mathbf{p}}^2$, otrzymujemy:

$$\hat{p}_{\mu}\hat{p}^{\mu}\zeta_{\pm} = 0, \quad (5.7)$$

dirfre

czyli równanie Kleina-Gordona dla cząstki o masie zero. Dla cząstek swobodnych można położyć

$$\zeta_{\pm}(x) = u_{\pm}(p)e^{-ipx}. \quad (5.8)$$

Podstawiając to wyrażenie do równania (5.7) i uwzględniając, że $i\partial_{\mu}e^{-ipx} = p_{\mu}e^{-ipx}$, otrzymujemy warunek na istnienie rozwiązań równań Weyla:

$$p^0 = \pm\sqrt{\mathbf{p}^2} = \pm|\mathbf{p}|. \quad (5.9)$$

Podobnie jak dla równania Kleina-Gordona, mamy więc rozwiązania z ujemną energią.

Podstawiając ζ_{\pm} ze wzoru (5.8) do równań Weyla otrzymujemy:

$$(p^0 \pm \mathbf{p}\boldsymbol{\sigma})u_{\pm}(p) = 0. \quad (5.10)$$

Dla cząstki o spinie $S = \frac{1}{2}$, operator skrętności

$$\lambda = \frac{\mathbf{p}\boldsymbol{\sigma}}{2|\mathbf{p}|}. \quad (5.11)$$

Dla $p^0 = |\mathbf{p}| > 0$ możemy więc, dzieląc obie strony każdego z równań Weyla przez $|\mathbf{p}|$, napisać

$$(1 \pm 2\lambda)u_{\pm} = 0. \quad (5.12)$$

Widać stąd, że równanie Weyla dla funkcji falowej $\xi^I = \zeta_-$ opisuje cząstki prawoskrętne ($\lambda = +\frac{1}{2}$), a równanie dla funkcji $\eta_j = \zeta_+$ cząstki lewoskrętne ($\lambda = -\frac{1}{2}$). Można to było przewidzieć, bo dla bardzo szybkich cząstek podwojona skrętność jest równa chiralności, a cząstka o masie zero i różnej od zera energii musi się poruszać z prędkością światła. Z teorii grupy obrotów dla cząstek z masą różną od zera wynika, że jeśli możliwa jest skrętność λ , to musi być też możliwa skrętność $-\lambda$. Cząstki Weyla mają jednak masę zero i dla nich nie ma takiej konieczności. Bardziej znany jest analogiczny przypadek fotonu, który ma stany $\lambda = \pm 1$, ale nie ma stanów z $\lambda = 0$. Znow jest to możliwe tylko dla tego, że masa fotonu jest równa zero. Foton wirtualny ma masę różną od zera, i co za tym idzie może być w stanach ze skrętnością $\lambda = 0$.

Dla $p^0 = -|\mathbf{p}|$ otrzymujemy równania

$$(1 \mp 2\lambda)u_{\pm} = 0. \quad (5.13)$$

To jest wynik dla niefizycznej cząstki o ujemnej energii. Przejście do antycząstek jednak, nie zmienia skrętności. Czy użyjemy interpretacji Stückelberga-Feynmana, czy Diraca zawsze wychodzi, że antycząstka ma w porównaniu do odpowiadającej jej cząstki o ujemnej energii przeciwny spin i przeciwny pęd. Rzut spinu na kierunek pędu nie ulega więc zmianie. Na przykład, przyjmując górne znaki, otrzymujemy lewoskrętną cząstkę i odpowiadającą jej prawoskrętną antycząstkę.

Doświadczalnie obserwuje się tylko lewoskrętne neutrino i prawoskrętne antyneutrino. Wybierając dla neutrino równanie Weyla na spinor ζ_+ , można by oba te wyniki wytłumaczyć, ale niestety tylko przy fałszywym założeniu, że masa neutrino jest równa zero. Obecnie tłumaczy się te fakty doświadczalne następująco. Funkcja falowa neutrino, podobnie jak elektronu, jest bispinorem. Z teorii słabych oddziaływań wynika, że we wzorach na amplitudy prawdopodobieństwa przejść ten bispinor jest zawsze mnożony lewostronnie przez czynnik $(1 - \gamma^5)$:

$$(1 - \gamma^5)\psi = 2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \xi^I \\ \eta_j \end{pmatrix} = 2 \begin{pmatrix} 0 \\ \eta_j \end{pmatrix}. \quad (5.14)$$

Tak więc aktywna jest tylko lewa składowa neutrino. Ponieważ neutrino poruszają się zwykle bardzo szybko, lewoskrętne neutrino ma dużą lewą składową, a prawoskrętne bardzo małą. Dlatego trudno jest nie tylko zaobserwować, ale nawet wytworzyć prawoskrętne neutrino. Funkcja falowa antyneutrino w teorii słabych oddziaływań jest zawsze mnożona przez czynnik $(1 + \gamma^5)$, co wyjaśnia dlaczego obserwowane antyneutrino są zawsze prawoskrętne.

5.2 Symetrie równań Weyla

Operacja \hat{O} jest symetrią równań Weyla, jeśli spinory $\hat{O}\zeta_{\pm}(x)$, spełniają te same równania, co spinory $\zeta_{\pm}(x)$. Jako symetrie wykluczone są więc transformacje, które zmieniają kropkowanie spinorów, bo spinory ζ_+ i ζ_- spełniają różne równania. Jak więc widać ze wzorów (4.85) i (4.86) ani inwersja przestrzenna, ani sprzężenie nie są symetrami równań Weyla. W szczególności antycząstka spełnia inne równanie niż cząstka. Jeśli transformacja nie zmienia kropkowania

spinora, to ponieważ dla danego rodzaju spinora jest tylko jedno równanie Weyla do wyboru, taka transformacja powinna być symetrią równania. Sprawdzimy to bezpośrednim rachunkiem dla transformacji T , CP i CPT .

Zacznijmy od odwrócenia czasu (4.94):

$$\hat{T}\xi^I(x) = \epsilon_T \xi_I^*(\hat{T}x) = \epsilon_T \epsilon \xi^{*I}(\hat{T}x), \quad \hat{T}\eta_I(x) = -\epsilon_T \eta^{*I}(\hat{T}x) = +\epsilon_T \epsilon \eta_I^*(\hat{T}x). \quad (5.15)$$

To daje

$$\hat{T}\zeta_{\pm}(x) = \epsilon_T \epsilon \zeta_{\pm}^*(\hat{T}x), \quad (5.16)$$

przy czym sprzężenie nie zmienia kropkowania wskaźników. Zwracamy uwagę, że ϵ bez wskaźnika oznacza tu macierz, a nie czynnik fazowy. Biorąc sprzężenie zespolone równania (5.6), otrzymujemy

$$(-\hat{p}^0 \mp \hat{p}_x \sigma_x \pm \hat{p}_y \sigma_y \mp \hat{p}_z \sigma_z) \zeta_{\pm}^*(x) = 0, \quad (5.17)$$

gdzie sprzężenie zespolone zmieniło znaki wszystkich operatorów \hat{p}^{μ} i macierzy σ_y . Zmieniając zmienną t na $-t$, zmieniamy znak operatora \hat{p}^0 i w argumentach funkcji falowej wprowadzamy $\hat{T}x$ w miejsce x :

$$(\hat{p}^0 \mp \hat{p}_x \sigma_x \pm \hat{p}_y \sigma_y \mp \hat{p}_z \sigma_z) \zeta_{\pm}^*(\hat{T}x) = 0. \quad (5.18)$$

Mnożąc to równanie lewostronnie przez macierz

$$\epsilon_T \epsilon = \epsilon_T i \sigma_y \quad (5.19)$$

i korzystając z reguł antykomutacji dla macierzy Pauli'ego, otrzymujemy

$$(\hat{p}^0 \pm \hat{p}_x \sigma_x \pm \hat{p}_y \sigma_y \pm \hat{p}_z \sigma_z) (\pm \epsilon_T \epsilon \zeta_{\pm}^*(\hat{T}x)) = (\hat{p}^0 \pm \mathbf{p}\boldsymbol{\sigma}) \hat{T}\zeta_{\pm}(x) = 0. \quad (5.20)$$

Tak więc odwrócenie czasu jest symetrią równań Weyla.

Podobnie ze wzorów (4.91):

$$\hat{C}\hat{P}\xi^I(x) = -i\xi_I^*(\hat{P}x) = -i\epsilon_C \epsilon \xi^{*I}(\hat{P}x), \quad \hat{C}\hat{P}\eta_I(x) = -i\eta^{*I}(\hat{P}x) = i\epsilon_C \epsilon \eta_I^*(\hat{P}x). \quad (5.21)$$

Biorąc sprzężenie zespolone równań Weyla i robiąc podstawienie $\mathbf{x} \rightarrow -\mathbf{x}$ dostajemy równanie

$$(-\hat{p}^0 \pm \hat{p}_x \sigma_x \mp \hat{p}_y \sigma_y \pm \hat{p}_z \sigma_z) \zeta_{\pm}^*(\hat{P}x) = 0. \quad (5.22)$$

Mnożąc obie strony tego równania lewostronnie przez macierze $\mp i\epsilon_c \epsilon$:

$$(-\hat{p}^0 \mp \hat{p}_x \sigma_x \mp \hat{p}_y \sigma_y \mp \hat{p}_z \sigma_z) (\mp i\epsilon_c \epsilon) \zeta_{\pm}^*(\hat{P}x) = -(\hat{p}^0 \pm \mathbf{p}\boldsymbol{\sigma}) \hat{C}\hat{P}\zeta_{\pm}(x) = 0, \quad (5.23)$$

a więc i transformacja CP jest symetrią równań Weyla. Z faktu, że inwersja przestrzenną nie jest symetrią równań Weyla wynika, że w świecie po drugiej stronie lustra funkcja falowa spełnia inne równania niż w naszym świecie. Ponieważ jednak transformacja CP jest dobrą symetrią równań Weyla, jednocześnie

przejście na drugą stronę lustra i zastąpienie cząstek przez ich antycząstki pozostawia te równania bez zmian.

Wreszcie dla przekształcenia CPT , ze wzorów (4.100)

$$\hat{C}\hat{P}\hat{T}\zeta_{\pm}(x) = \mp i\epsilon_C\epsilon_T^*\zeta_{\pm}(-x). \quad (5.24)$$

Pomnożenie przez stałe $\mp i\epsilon_C\epsilon_T^*$ i zamiana x na $-x$ nie zmienia równań Weyla, więc mamy

$$(\hat{p}^0 \pm \mathbf{p}\boldsymbol{\sigma})\hat{C}\hat{P}\hat{T}\zeta_{\pm}(x) = 0. \quad (5.25)$$

Ta symetria wynika też z uwagi, że operacje \hat{T} i $\hat{C}\hat{P}$ są symetriami teorii.

5.3 Równanie Diraca

Równanie Diraca opisuje cząstkę o masie m i spinie $\frac{1}{2}$, przy czym zakłada się, że przestrzeń rozwiązań jest niezmiennicza względem rozszerzonej grupy Lorentza, to znaczy względem właściwych transformacji Lorentza i inwersji przestrzennej. Jak zawsze dla cząstek swobodnych, przyjmujemy też niezmienniczość tej przestrzeni względem translacji. Wynika stąd, że funkcja falowa cząstki swobodnej zależy od punktochwili tylko poprzez czynnik e^{-ipx} . Zakładamy, że rozwiązania równania Diraca spełniają także równanie Kleina-Gordona, więc składowe czteropędu spełniają relację

$$p^0 = \pm\sqrt{\mathbf{p}^2 + m^2}. \quad (5.26)$$

Do budowy funkcji falowej mamy teraz dwa rodzaje spinorów: $\xi^I(x)$ i $\eta_I(x) = -i\hat{P}\xi^I(\hat{P}x)$. W dalszym ciągu nie wypisujemy już wskaźników spinorowych.

Używając spinorów kropkowanych i niekropkowanych, możemy zastąpić dwa niezależne równania Weyla przez jeden kowariantny układ równań

$$\begin{aligned} (p_0 + \boldsymbol{\sigma}\hat{\mathbf{p}})\xi(x) &= m_1\eta(x), \\ (p_0 - \boldsymbol{\sigma}\hat{\mathbf{p}})\eta(x) &= m_2\xi(x), \end{aligned} \quad (5.27)$$

gdzie m_1 i m_2 są stałymi o wymiarze masy. W układzie spoczynkowym, dla cząstki o dodatniej energii, mamy $p_0 = m$, $\mathbf{p} = \mathbf{0}$ i $\xi(x) = \eta(x)$. Podstawiając do powyższych równań, otrzymujemy

$$m_1 = m_2 = m. \quad (5.28)$$

Mnożąc pierwsze z równań (5.27) obustronnie przez operator $(p_0 - \boldsymbol{\sigma}\hat{\mathbf{p}})$ i korzystając z równania Kleina-Gordona, otrzymujemy drugie. To znaczy, że z czterech składowych spinorów ξ i η tylko dwie są niezależne. Dwie, na przykład spinor η lub spinor ξ , można wybrać dowolnie i użyć równań (5.27) do wyznaczenia dwu pozostałych.

Korzystając z definicji macierzy γ^μ , możemy zapisać oba równania (5.27) jako jedno równanie na bispinor

$$(\not{p} - m)\psi(x) = 0. \quad (5.29)$$

To jest słynne **równanie Diraca** (1928). Operator $(\hat{p} - m)$ nazywa się **operatorem Diraca**.

Ćwiczenie Pokazać, że działając operatorem \hat{P} na obie strony pierwszego z równań (5.27) otrzymuje się drugie, jeśli $m_1 = m_2$. Wyjaśnić, dlaczego wybrany został argument wykorzystujący równanie Kleina-Gordona.

•••••

Ćwiczenie Pokazać, że dwukomponentowe równanie

$$\hat{P}^{A\dot{B}}\hat{P}_{\dot{B}C}\xi^C = m^2\xi^A$$

jest równoważne z równaniem Diraca. Cztery komponenty w równaniu Diraca są potrzebne do opisu dwu stanów spinowych elektronu i dwu stanów spinowych pozytonu. Wyjaśnić, dlaczego w zaproponowanym równaniu wystarczą do tego celu dwie komponenty.

Wskazówki: Zdefiniować pomocniczy spinor $\eta_{\dot{B}} = \frac{1}{m}\hat{P}_{\dot{B}C}\xi^C$. Zauważyć, że równanie jest drugiego rzędu względem czasu.

•••••

Wprowadzając pole elektromagnetyczne sprzężone minimalnie, otrzymujemy równanie

$$(\hat{p} - q\hat{A} - m)\psi(x) = 0, \quad (5.30)$$

gdzie q jest ładunkiem elektrycznym cząstki. Dla elektronu o energii $E \gg m \approx 0.5\text{MeV}$ można w dobrym przybliżeniu zaniedbać człon proporcjonalny do masy. Wtedy równanie rozpada się na dwa niezależne równania Weyla: jedno dla elektronów prawoskrętnych, drugie dla lewoskrętnych. Wynika stąd w szczególności, że skrętność elektronu o wysokiej energii, jeśli jest określona w stanie początkowym, nie ulega zmianie przy rozpraszaniu w polu elektromagnetycznym.

Ćwiczenie Elektron o wysokiej energii, który ma określoną skrętność początkową, krąży po obwodzie koła w jednorodnym polu magnetycznym. Pokazać, że według równania Diraca okres precesji spinu jest taki sam jak okres obiegu elektronu. Doświadczalnie stwierdza się niewielkie różnice między tymi dwoma okresami. Jaką wielkość można wyznaczyć mierząc te różnice?

•••••

Wyprowadzimy teraz równoważne równanie na funkcję $\bar{\psi}(x)$. W tym celu bierzemy sprzężenie hermitowskie poprzedniego równania i, pamiętając, że macierz γ^0 jest hermitowska, a pozostałe macierze γ^μ antyhermitowskie, otrzymujemy:

$$\psi^\dagger(x)(-\hat{p}_0\gamma^0 - \mathbf{p}\boldsymbol{\gamma} - qA_0\gamma^0 - q\mathbf{A}\boldsymbol{\gamma} - m) = 0 \quad (5.31)$$

Zauważmy, że jest to sprzężenie hermitowskie w przestrzeni macierzy 4×4 wymiarowych, przy czym $i \rightarrow -i$, a nie w przestrzeni Hilberta, gdzie byłoby $\hat{p}_j^\dagger = \hat{p}_j$. Operatory różniczkowania działają na funkcję $\psi^\dagger(x)$. Mnożąc powyższe równanie prawostronnie przez $-\gamma^0$ otrzymujemy

$$\bar{\psi}(x)(\overleftarrow{\hat{p}} + q\hat{A} + m) = 0. \quad (5.32)$$

duadir

Skierowana w lewo strzałka nad operatorem przypomina, że operator działa na funkcję w lewo od niego. To równanie bywa nazywane równaniem dualnym.

Ćwiczenie Pogodzić ze sobą następujące dwie obserwacje. Operator $-i\frac{\partial}{\partial x}$ może być rozpatrywany jako macierz 1×1 wymiarowa. Dla takiej macierzy sprzężenie hermitowskie jest identyczne ze zwykłym sprzężeniem zespolonym, i w tym przypadku powoduje zmianę znaku. Z drugiej strony wiadomo, że operator $\hat{p}_x = -i\frac{\partial}{\partial x}$ jest hermitowski, a zatem nie zmienia się przy sprzężeniu hermitowskim.

•••••

Sprawdźmy teraz, że równanie Diraca, ze sprzężonym minimalnie potencjałem wektorowym, daje prąd zachowywany. Mnożąc równanie Diraca lewostronnie przez $\bar{\psi}(x)$ i dodając do niego równanie dualne pomnożone prawostronnie przez $\psi(x)$ otrzymujemy:

$$\bar{\psi}(x) \overleftarrow{\not{p}} \psi(x) + \bar{\psi}(x) \not{p} \psi(x) = i\partial_\mu [\bar{\psi}(x) \gamma^\mu \psi(x)] = 0. \quad (5.33)$$

Stąd prąd zachowywany ma postać¹

$$J^\mu = N(p) \bar{\psi}(x) \gamma^\mu \psi(x), \quad (5.34)$$

gdzie $N(p)$ jest liczbowym współczynnikiem normalizacyjnym. Zauważmy, że na mocy dyskusji z paragrafu 4.12 ta wielkość jest czterowektorem i że wśród kowariantów dwuliniowych nie było innego czterowektora do wyboru, ta ostatnia uwaga nie stanowi oczywiście dowodu, że ten prąd jest zachowywany. W szczególności, dla zerowej składowej prądu otrzymujemy

$$\bar{\psi}(x) \gamma^0 \psi(x) = \psi^\dagger(x) \psi(x) \geq 0. \quad (5.35)$$

Dirac, który odkrył ten fakt w czasach, kiedy składowa prądu J^0 była powszechnie uważana za gęstość prawdopodobieństwa, uznał że to jest wielki sukces teorii. Później, kiedy J^0 zostało zinterpretowane jako gęstość ładunku, ten wynik stał się kłopotliwy, bo gęstości ładunku dla elektronów i pozytonów powinny mieć przeciwne znaki. Rozwiązanie tej trudności podamy w następnym paragrafie.

Pokażemy jeszcze, że równanie Diraca może być zapisane w schrödingerowskiej postaci

$$i \frac{\partial \psi(x)}{\partial t} = \hat{H} \psi(x). \quad (5.36)$$

W tym celu w równaniu (5.29) przenosimy na prawą stronę wszystkie wyrazy oprócz zawierającego różniczkowanie po czasie. Otrzymujemy

$$i\gamma^0 \frac{\partial \psi(x)}{\partial t} = (\boldsymbol{\gamma} \hat{\mathbf{p}} + m) \psi(x). \quad (5.37)$$

Następnie, mnożąc lewostronnie obie strony równania przez macierz γ^0 i porównując wynik z równaniem Schrödingera mamy

$$\hat{H} = \boldsymbol{\alpha} \hat{\mathbf{p}} + \beta m, \quad (5.38)$$

gdzie wprowadziliśmy oznaczenia Diraca

¹Ten sam wynik otrzymalibyśmy, gdyby do masy cząstki dodać jakikolwiek rzeczywisty skalarny potencjał $V(x)$.

$$\alpha = \gamma^0 \gamma, \quad \beta = \gamma^0. \quad (5.39)$$

alfadi

Macierze α_i można zapisać w postaci

$$\alpha_i = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} = \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix}. \quad (5.40)$$

Są to więc macierze hermitowskie, jak potrzeba żeby hermitowski był hamiltonian. Omówimy jeszcze krótko dwa wyniki dotyczące równania Diraca.

Każde rozwiązanie równania Diraca, podobnie jak nierelatywistyczna funkcja falowa, może być przedstawione jako superpozycja fal płaskich. Jeśli w tej superpozycji występują zarówno człony z $p_0 > m$, jak i człony z $p_0 < -m$, to we wzorach na wartości średnie operatorów pojawiają się człony interferencyjne, oscylujące z częstością większą niż $2m$. Podstawiając masę elektronu otrzymujemy częstość około 10^{21} Hz. W szczególności, w odniesieniu do operatora położenia wynika stąd, że cząstka wykonuje gwałtowne, krótkozasięgowe ruchy w różnych kierunkach, które dopiero po wyśredniowaniu dają makroskopowy ruch. Dla tego zjawiska przyjęła się niemiecka nazwa **Zitterbewegung** (Schrödinger 1930) – ruch polegający na drżeniu, dygotaniu czy trzęsieniu się. Dla elektronu duża częstość tych ruchów uniemożliwia ich obserwację, ale w innych działach fizyki (fizyka fazy skondensowanej, kondensaty Einsteina) są układy, do których stosuje się równanie Diraca, choć z zupełnie inną interpretacją fizyczną, i tam Zitterbewegung można zobaczyć.

Dla elektronu blisko granicy nierelatywistycznej, wkład od członów z $p_0 < -m$ można zaniedbać, ale człony z $p_0 > m$ nie stanowią układu zupełnego, więc nie każdy rozkład elektronu w przestrzeni da się z ich pomocą zrealizować. W szczególności, **Newton i Wigner** wykazali², że z takich fal nie da się zbudować pakietu falowego o promieniu dużo mniejszym od $\frac{1}{m}$. Podstawiając masę elektronu, otrzymujemy promień minimalny około 4×10^{-11} cm. To znaczy, że prawie nierelatywistyczny elektron należy sobie wyobrazić nie jako cząstkę punktową, ale jako chmurkę o promieniu rzędu m_e^{-1} . Ta chmurka jest mała w stosunku do promienia atomu (rzędu 10^{-8} cm), ale duża w stosunku do promienia protonu (rzędu 10^{-13} cm). Jak zobaczymy w paragrafie 5.8, to rozmycie elektronu wpływa na poziomy energetyczne atomu wodoru. Muon jest około 200 razy cięższy od elektronu. To znaczy, że dla niego minimalny promień rozmycia jest rzędu 10^{-13} cm – porównywalny z promieniem protonu. Dzięki temu, badając spektroskopię atomu wodoru z elektronem zastąpionym przez muon, można uzyskać informacje o promieniu ładunkowym protonu.

5.4 Symetrie równania Diraca

Podobnie jak dla równań Weyla, kowariancja równania Diraca względem właściwych transformacji Lorentza wynika ze sposobu w jaki to równanie zostało skonstruowane. Pozostają do zbadania transformacje $O = P, C, T$. Operacja O jest symetrią równania, lub równoważnie równanie jest kowariantne względem przekształcenia O , jeśli funkcja $\psi^O(x)$ spełnia to samo równanie, co funkcja $\psi(x)$. Funkcje $\psi^O(x)$ wyliczamy jak w paragrafie 4.10.

²Ich dowód jest krytykowany z punktu widzenia kwantowej teorii pola, ale wynik uchodzi za poprawny.

Punktem wyjścia w dowodach symetrii jest zawsze równanie Diraca

$$(\not{p} - m)\psi(x) = 0. \quad (5.41)$$

Zmieniając zmienne z \mathbf{x} na $-\mathbf{x}$ i mnożąc lewostronnie przez macierz $U_P = i\gamma^0$, otrzymujemy

$$U_P(\hat{p}^0\gamma^0 + \hat{\mathbf{p}}\boldsymbol{\gamma} - m)\psi(\hat{P}x) = 0. \quad (5.42)$$

Stąd, korzystając z reguł antykomutacji dla macierzy $\boldsymbol{\gamma}$,

$$(\not{p} - m)\psi^P(x) = 0. \quad (5.43)$$

Dla cząstki swobodnej inwersja przestrzenna jest więc symetrią równania Diraca. Wprowadzenie oddziaływań może tę symetrię zachować, lub złamać. Zwykle rozpatruje się przypadek, kiedy inwersji przestrzennej ulega nie tylko cząstka, ale i wszystkie pola, które na nią działają. Tak rozumiana, symetria jest zachowywana przez oddziaływania silne i elektromagnetyczne, jest natomiast łamana przez oddziaływania słabe.

Znaczenie sprzężenia w relatywistycznej mechanice kwantowej wynika głównie stąd, że bispinory ψ^C odpowiadające rozwiązaniom równania Diraca ψ z ujemnymi energiami opisują obserwowane doświadczalnie antycząstki. W porównaniu do cząstek opisywanych przez spinory ψ z dodatnimi energiami, te antycząstki mają takie same spin i masę, natomiast przeciwny ładunek – tak jak powinno być. Odtąd, zgodnie z powszechnie przyjmowaną przez fizyków terminologią, zamiast sprzężenie będziemy pisać sprzężenie ładunkowe.

Biorąc sprzężenie zespolone równania Diraca i mnożąc je lewostronnie przez macierz $U_C = i\gamma^y$ otrzymujemy równanie

dcomco

$$U_C(-\hat{p}^0\gamma^0 + \hat{p}_x\gamma^x - \hat{p}_y\gamma^y + \hat{p}_z\gamma^z - m)\psi^*(x) = 0. \quad (5.44)$$

Przy sprzężeniu zespolonym zmieniały znaki wszystkie operatory $\hat{p}^\mu = i\partial^\mu$ i czysto urojona macierz $\boldsymbol{\gamma}^y$. Korzystając z reguł antykomutacji dla macierzy $\boldsymbol{\gamma}^\mu$ otrzymujemy

$$(\not{p} - m)\psi^C(x) = 0. \quad (5.45)$$

Sprzężenie ładunkowe jest więc symetrią równania Diraca dla cząstki swobodnej. Stosując to samo rozumowanie do funkcji falowej cząstki sprzężonej minimalnie z polem elektromagnetycznym, a więc spełniającej równanie

$$(\not{p} - q\mathcal{A} - m)\psi(x) = 0, \quad (5.46)$$

otrzymujemy

$$(\not{p} + q\mathcal{A} - m)\psi^C(x) = 0, \quad (5.47)$$

Zmiana znaku członu $q\mathcal{A}$ bierze się stąd, że składowe czterowektora $A^\mu(x)$, w przeciwieństwie do operatorów \hat{p}^μ , nie zmieniają znaku przy sprzężeniu zespolonym. To równanie jest równaniem Diraca dla cząstki o ładunku elektrycznym $-q$ i w tym sensie symetria jest złamana. Jeśli jednak, analogicznie jak dla inwersji, przyjmiemy że sprzężeniu ładunkowemu podlega nie tylko cząstka, ale i wszystkie źródła pola, to pole A też zmienia znak i mamy symetrię względem

sprężenia ładunkowego. Tak rozumiana symetria obowiązuje dla oddziaływań silnych i elektromagnetycznych, ale jest łamana dla oddziaływań słabych.

Ćwiczenie Pokazać, że w reprezentacji spinorowej, dla $\mu = 0, 1, 2, 3$, zachodzi: $\gamma^y \gamma^{\mu*} = -\gamma^\mu \gamma^y$. Wykorzystać tę tożsamość dla wyprowadzenia powyższego równania na funkcję ψ^C .

•••••

W poprzednim paragrafie zwróciliśmy uwagę, że zerowa składowa prądu zachowywanego,

$$J^0 = N\psi^\dagger\psi = N\sum_{i=1}^4\psi_i^*\psi_i, \quad (5.48)$$

ma znak niezależny od tego, czy energia jest dodatnia czy ujemna. Jest to sprzeczne z interpretowaniem J^0 jako gęstości ładunku. Nie poprawia sytuacji rozważanie prądu dla antycząstki:

$$J_C^0 = N\psi^{C\dagger}\psi^C = N\sum_{i=1}^4\psi_i^{*C}\psi_i^C, \quad (5.49)$$

bo i tu zerowa składowa ma stały znak. Podstawiając $\psi^C = i\gamma^y\psi^*$ i $\psi^{C\dagger} = -i\psi^T\gamma^{y\dagger} = +i\psi^T\gamma^y$ otrzymujemy

$$J_C^0 = Ni\psi^T\gamma^y i\gamma^y\psi^* = N\psi^T\psi^* = N\sum_i\psi_i\psi_i^*. \quad (5.50)$$

W mechanice kwantowej, gdzie funkcje falowe są zwykłymi liczbami, ten wzór jest równoważny ze wzorem (5.48), ale w kwantowej teorii pola wielkości ψ i ψ^\dagger są operatorami pola, które w przypadku fermionów antykomutują między sobą. Przy takiej interpretacji, różnica w kolejności czynników we wzorach dla cząstek i antycząstek wprowadza różnicę w znaku, i gęstość ładunku antycząstek ma znak przeciwny niż gęstość ładunku cząstek tak jak powinno być. W mechanice kwantowej możemy się umówić, że dla antycząstek wstawimy zamiast stałej N stałą $N_C = -N$, ale nie da się tego uzasadnić jako wniosku z symetrii. Uzasadnienie zmiany znaku J_0 przy przejściu od cząstek do antycząstek istnieje więc, ale dopiero na gruncie kwantowej teorii pola.

Pokażemy jeszcze, że parzystość wewnętrzna układu cząstka-antycząstka jest minus. Żeby wyeliminować parzystość orbitalną, rozpatrujemy cząstkę i jej antycząstkę, obie w spoczynku. W tej sytuacji równanie Diraca dla cząstki ($p^0 > 0$) przybiera, po wyeliminowaniu czynnika e^{-ipx} , postać

$$p^0\gamma^0u(\mathbf{0}) = mu(\mathbf{0}). \quad (5.51)$$

W zapisie wykorzystaliśmy tu fakt, że jeśli znak składowej pędu p^0 jest znany, to bispinor u zależy tylko od przestrzennych składowych pędu. W układzie spoczynkowym $p^0 = m$, więc

$$u(\mathbf{0}) = \begin{pmatrix} \xi \\ \xi \end{pmatrix}. \quad (5.52)$$

Pod działaniem operatora inwersji przestrzennej

$$\hat{P}u(\mathbf{0}) = i\gamma^0u(\mathbf{0}) = iu(\mathbf{0}). \quad (5.53)$$

Dla antycząstki $p_0 = -m$ i bispinor $u^C(\mathbf{0})$ jest rozwiązaniem tego samego równania, też z dodatnią energią, więc operator \hat{P} działa na niego tak samo jak na bispinor $u(\mathbf{0})$, i dla pary

$$\hat{P}(u(\mathbf{0})u^C(\mathbf{0})) = iu(\mathbf{0})iu^C(\mathbf{0}) = -(u(\mathbf{0})u^C(\mathbf{0})). \quad (5.54)$$

Parzystość wewnętrzna pary cząstka i jej antycząstka jest ujemna.

Parzystość wewnętrzna pary cząstka-antycząstka jest iloczynem parzystości wewnętrznych cząstki i antycząstki: $P_a P_{\bar{a}} = -1$. Jeśli przyjmiemy $P_a^2 = P_{\bar{a}}^2 = +1$, to widać, że $P_a = -P_{\bar{a}}$, a więc cząstka i antycząstka nie mogą być identyczne. Jeśli natomiast $P_a^2 = P_{\bar{a}}^2 = -1$, to mamy $P_a = P_{\bar{a}} = \pm i$, co nie przeczy założeniu $a = \bar{a}$. Wobec popularności hipotezy, że neutrino i antyneutrino są identyczne, przyjmuje się obecnie, że w działaniu na spinory $\hat{P}^2 = -1$. Zauważmy, że chociaż znaki parzystości cząstki i antycząstki są kwestią konwencji, bo można je zmienić robiąc nieobserwowalny obrót układu współrzędnych o kąt 2π wokół dowolnej osi, to ich względny znak jest dobrze określony.

Biorąc sprzężenie zespolone równania Diraca, zmieniając zmienną t na $-t$ i mnożąc lewostronnie wynik przez macierz $U_T = \gamma^x \gamma^z$ otrzymujemy równanie

$$U_T(\hat{p}^0 \gamma^0 + \hat{p}_x \gamma^x - \hat{p}_y \gamma^y + \hat{p}_z \gamma^z - m)\psi^*(\hat{T}x) = 0, \quad (5.55)$$

co po zastosowaniu reguł antykomutacji dla macierzy γ daje

$$(\hat{p} - m)\psi^T(x) = 0. \quad (5.56)$$

Tak więc i odwrócenie czasu jest symetrią równania Diraca dla cząstki swobodnej. Ta symetria pozostaje w mocy, jeśli włączymy oddziaływania silne i elektromagnetyczne, oczywiście przy zwykłym warunku, że odwracane w czasie są też i pola działające na cząstkę. Przy oddziaływaniach słabych ta symetria jest łamana, chociaż znacznie słabej niż symetrie C i P .

Z kowariancji względem odbić C, P i T , wynika kowariancja względem ich złożenia, w szczególności względem transformacji CP i CPT . Jak dyskutowaliśmy omawiając symetrie równania Kleina-Gordona, symetria CPT obowiązuje dla bardzo szerokiej klasy teorii. Dlatego przyjmuje się, że symetria CP jest łamana wtedy i tylko wtedy, kiedy jest łamana symetria względem odwrócenia czasu.

Ćwiczenie Sprawdź, że gdyby we wzorach na odwrócenie czasu w spinorach kropkowanych zmienić znak wyniku, to otrzymałoby się $U_{CPT} = i$, i operacja CPT nie byłaby symetrią równania Diraca. Przedyskutować ten wynik.

•••••

Ćwiczenie Sprawdzić, że gdyby pod działaniem sprzężenia na spinory niekropwane powstawały spinory kropkowane różniące się od spinorów kropkowanych otrzymanych pod działaniem inwersji przestrzennej bardziej niż o stały czynnik, to operacja CPT nie byłaby symetrią równania Diraca. Przedyskutować ten wynik.

•••••

5.5 Cząstka swobodna – reprezentacja spinorowa

Przedyskutujemy teraz, używając reprezentacji spinorowej, rozwiązania równania Diraca dla cząstki swobodnej. Zgodnie z ogólną teorią, będziemy szukali

rozwiązań w postaci

$$\psi(x) = u(p)e^{-ipx}, \quad u(p) = \begin{pmatrix} \xi(p) \\ \eta(p) \end{pmatrix}, \quad (5.57)$$

gdzie czterowektor p jest czteropędem cząstki, na którego składowe nałożony jest warunek (5.26). Podstawiając te wyrażenia do równania Diraca, otrzymujemy równania na spinory:

$$(p_0 + \boldsymbol{\sigma}\mathbf{p})\eta = m\xi, \quad (p_0 - \boldsymbol{\sigma}\mathbf{p})\xi = m\eta. \quad (5.58)$$

Ćwiczenie Pokazać, że eliminując z równań Diraca spinor ξ otrzymuje się równanie Kleina Gordona na spinor η .

•••••

Spinor ξ możemy wybrać dowolnie. Wygodnie jest przyjąć normalizację

$$\xi^\dagger \xi = 1. \quad (5.59)$$

Spinor η można wyrazić przez spinor ξ przy pomocy drugiego z równań (5.58). Otrzymujemy

$$u_\pm(p) = N_\pm \begin{pmatrix} \xi \\ \frac{p_0 - \boldsymbol{\sigma}\mathbf{p}}{m} \xi \end{pmatrix}, \quad (5.60)$$

gdzie wskaźniki \pm określają znak p_0 .

Wygodnie jest wybrać, jako dwa niezależne spinory ξ , wektory własne operatora skrętności

$$\boldsymbol{\sigma}\mathbf{p}\xi_\lambda = 2\lambda|\mathbf{p}|\xi_\lambda, \quad \lambda = \pm\frac{1}{2}. \quad (5.61)$$

Mamy więc cztery rozwiązania

$$u_{\lambda\pm}(p) = N_\pm \begin{pmatrix} \xi_\lambda \\ \frac{p^0 - 2\lambda|\mathbf{p}|}{m} \xi_\lambda \end{pmatrix}, \quad \lambda = \pm\frac{1}{2}, \quad \pm p_0 = E > 0. \quad (5.62)$$

Możemy podstawić

$$p_0 = \pm m \cosh y, \quad |\mathbf{p}| = m \sinh y, \quad (5.63)$$

i otrzymujemy

$$u_{\lambda\pm}(p) = N_\pm \begin{pmatrix} \xi_\lambda \\ \pm e^{\mp 2\lambda y} \xi_\lambda \end{pmatrix}. \quad (5.64)$$

Widać, że przy $p_0 > 0$ i wysokich energiach, kiedy $y \gg 1$, górny spinor dominuje przy $\lambda = +\frac{1}{2}$, a dolny przy $\lambda = -\frac{1}{2}$. Przy wysokich energiach to upraszcza rachunki i powoduje, że reprezentacja spinorowa jest wtedy wygodna. Natomiast przy niskich energiach, kiedy $|y| \lesssim 1$, oba spinory są w reprezentacji spinorowej porównywalne i wygodniej jest używać reprezentacji Diraca, którą omówimy w następnym paragrafie.

Ćwiczenie Pokazać, że przy wysokich energiach prawdopodobieństwo tego, że cząstka prawa okaże się lewoskrętna, czy że cząstka lewoskrętna okaże się prawa, jest rzędu $\frac{m^2}{2p_0^2}$.

•••••

Warunek normalizacji nałożymy na skalarne wielkości

$$\bar{\psi}_{\pm}\psi_{\pm} = \bar{u}_{\pm}(p)u_{\pm}(p) = \pm 2|N_{\pm}|^2 e^{\mp 2\lambda y}. \quad (5.65)$$

Widać, że normalizować trzeba do liczby o takim znaku, jak znak p_0 . Przyjmując liczbę dwa, otrzymujemy

$$N_{\pm} = e^{\pm\lambda y}, \quad (5.66)$$

gdzie, jak zwykle, wybraliśmy stałą normalizacyjną rzeczywistą i dodatnią.

Możemy teraz wypisać wszystkie cztery rozwiązania na funkcje falowe $\psi(x)$: *pzdm*od

$$\psi_{\lambda\pm}(x) = \begin{pmatrix} e^{\pm\lambda y}\xi_{\lambda} \\ \pm e^{\mp\lambda y}\xi_{\lambda} \end{pmatrix} e^{-ipx}. \quad (5.67)$$

Ćwiczenie Rozwiązując równanie Diraca pokazać, że dla cząstki ($p_0 > 0$) w spoczynku $u_+(p) = \begin{pmatrix} \xi_{\lambda} \\ \xi_{\lambda} \end{pmatrix}$ i przez odpowiednie pchnięcie uzyskać z tego rozwiązania rozwiązanie (5.67), znormalizowane zgodnie z powyższą konwencją. Zauważyć, że wynika to z relatywistycznej kowariancji równania Diraca i przyjętej normalizacji.

•••••

Prąd zachowywany dla $p_0 > 0$ najłatwiej jest wyliczyć w dwu krokach. W układzie spoczynkowym cząstki $\mathbf{p} = \mathbf{0}$, $y = 0$ i

$$u(p) = \begin{pmatrix} \xi_{\lambda} \\ \xi_{\lambda} \end{pmatrix}, \quad J^0 = u^{\dagger}u = 2, \quad (5.68)$$

gdzie pominęliśmy nieistotny dla naszej dyskusji stały czynnik w definicji prądu. Składowe przestrzenne tego prądu $\mathbf{J} = \mathbf{0}$, bo żaden kierunek w przestrzeni nie jest wyróżniony. Ponieważ J^{μ} jest czterowektorem, można go jednoznacznie określić jako ten czterowektor, który w układzie spoczynkowym ma składowe $\{2, \mathbf{0}\}$. Tym czterowektorem jest

$$J^{\mu} = 2\frac{p^{\mu}}{m}. \quad (5.69)$$

Ćwiczenie Znaleźć składowe prądu J^{μ} wprost z definicji, nie używając podanego w tekście skrótu.

•••••

Żeby uzyskać rozwiązania dla pozytonów, trzeba dokonać sprzężenia ładunkowego rozwiązań z $p_0 < 0$

$$\psi_{\lambda}^C(x) \equiv u_{\lambda}^C(p_C)e^{+ipx} = \begin{pmatrix} 0 & -\epsilon \\ \epsilon & 0 \end{pmatrix} u_{\lambda}^*(p)e^{-ip_Cx}, \quad (5.70)$$

gdzie zastosowaliśmy oznaczenie $p_C = -p$. Antycząstka ma pęd przeciwny do pędu cząstki, a skrętność taką jak cząstka. W powyższym wzorze zależność

od czasu i współrzędnych jest taka jak dla cząstki o pędzie p_C . Pozostaje do sprawdzenia skrętność. Wykonując mnożenie otrzymujemy

$$u_\lambda^C(p_C) = \begin{pmatrix} 0 & -\epsilon \\ \epsilon & 0 \end{pmatrix} u_\lambda^*(p) = \begin{pmatrix} e^{\lambda y} \epsilon \xi_\lambda^* \\ e^{-\lambda y} \epsilon \xi_\lambda^* \end{pmatrix}. \quad (5.71)$$

epsist

Żeby zinterpretować spinory występujące w tym wyrażeniu przyda się tożsamość

$$\epsilon \sigma_i^* = i \sigma_y \sigma_i^* = -\sigma_i \epsilon, \quad i = x, y, z. \quad (5.72)$$

Biorąc sprzężenie zespolone równości (5.61), mnożąc obie strony lewostronnie przez macierz $-\epsilon$ i korzystając z powyższej tożsamości dostajemy

$$\sigma \mathbf{p} \epsilon \xi_\lambda^* = -2\lambda |\mathbf{p}| \epsilon \xi_\lambda^*, \quad (5.73)$$

lub równoważnie

$$\sigma \mathbf{p}_C \epsilon \xi_\lambda^* = 2\lambda |\mathbf{p}_C| \epsilon \xi_\lambda^*. \quad (5.74)$$

Spinor $\epsilon \xi_\lambda^*$ opisuje więc cząstkę o rzucie spinu na kierunek \mathbf{p}_C równym λ i można wprowadzić oznaczenie

$$\epsilon \xi_\lambda^* = \xi_\lambda^C. \quad (5.75)$$

Mamy więc ostatecznie

$$\psi_\lambda^C(x) = \begin{pmatrix} e^{\lambda y} \xi_\lambda^C \\ e^{-\lambda y} \xi_\lambda^C \end{pmatrix} e^{-ip_C x}, \quad (5.76)$$

czyli rozwiązanie różniące się od rozwiązania dla $p_0 > 0$ tylko zmianą oznaczeń z p na p_C i ξ na ξ^C . Widać tu pełną symetrię między cząstkami i ich antycząstkami. Można by równie dobrze nazwać pozytony cząstkami, przypisać im funkcje falowe $\psi^C(x)$ i uzyskać funkcje falowe dla elektronów przez sprzężenie ładunkowe. Warto jeszcze wiedzieć, że w praktycznych rachunkach często wygodniej jest operować cząstkami z ujemną energią niż odpowiadającymi im antycząstkami.

Innym problemem, dla którego równanie Diraca da się dokładnie rozwiązać, jest ruch elektronu poruszającego się w polu kulombowskim punktowego, nieskończenie ciężkiego ładunku. Można wyliczyć amplitudy rozpraszania oraz energie i funkcje falowe stanów związanych. Do tych stanów związanych wrócimy w paragrafach 5.9 i 5.10, ale będziemy stosowali metody przybliżone, których wyniki są znacznie bardziej pouczające niż wyniki dokładne.

5.6 Cząstka swobodna – reprezentacja Diraca

Reprezentacja Diraca jest często używana w praktyce. Jest szczególnie dogodna, kiedy jesteśmy blisko granicy nierelatywistycznej i poprawki relatywistyczne są małe. Macierze gamma w reprezentacji Diraca, γ_D^μ , są związane z macierzami gamma w reprezentacji spinorowej γ_S^μ wzorem

$$\gamma_D^\mu = U \gamma_S^\mu U^\dagger, \quad (5.77)$$

uspidi

gdzie macierz unitarna

$$U = U^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(\gamma_S^0 + \gamma_S^5). \quad (5.78)$$

Ćwiczenie Pokazać, że w reprezentacji Diraca dolna składowa bispinora jest proporcjonalna do $\xi^I(x) - \eta_j(x)$ i zeruje się dla cząstki swobodnej w spoczynku. Jakie zastosowania może mieć ten wynik?

•••••

Mnożąc macierze, lub prościej korzystając z własności komutacyjnych macierzy gamma, otrzymujemy

$$\gamma_D^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \gamma_S^5, \quad (5.79)$$

$$\gamma_D = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix} = -\gamma_S, \quad (5.80)$$

$$\gamma_D^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \gamma_S^0. \quad (5.81)$$

Żeby rozwiązać w reprezentacji Diraca równanie Diraca dla cząstki swobodnej, robimy zwykle podstawienie w funkcji falowej

$$\psi(x) = N \begin{pmatrix} \phi(p) \\ \chi(p) \end{pmatrix} e^{-ipx}. \quad (5.82)$$

Wielkości ϕ i χ , są spinorami względem grupy obrotów, ale przy transformacjach Lorentza transformują się jak superpozycje spinorów kropkowanych i niekropkowanych. Mimo to, dla wygody, będziemy je nazywali spinorami. Podstawiając do równania Diraca, otrzymujemy równania macierzowe na spinory ϕ i χ : dirdir

$$(p_0 - m)\phi = \boldsymbol{\sigma}\mathbf{p}\chi, \quad (p_0 + m)\chi = \boldsymbol{\sigma}\mathbf{p}\phi. \quad (5.83)$$

Rozpatrzymy najpierw przypadek

$$p_0 = E > 0. \quad (5.84)$$

Rozwiązanie równania Diraca można uzyskać, wybierając dowolnie spinor ϕ i wyliczając odpowiadający mu spinor χ . Wybieramy spinory ϕ_s spełniające równości defxil

$$\sigma_z \phi_s = 2s \phi_s, \quad \phi_s^\dagger \phi_s = 1. \quad (5.85)$$

Jak widać z tego równania, $s = \pm \frac{1}{2}$ jest rzutem spinu cząstki opisywanej przez spinor ϕ_s na oś z . Używając drugiego z równań (5.83), wyrażamy spinor χ_s przez spinor ϕ_s i otrzymujemy dwa niezależne rozwiązania dirpoz

$$\psi_{+,s}(x) = N_+ \begin{pmatrix} \phi_s \\ \frac{\boldsymbol{\sigma}\mathbf{p}}{E+m} \phi_s \end{pmatrix} e^{-ipx}, \quad p_0 > 0, \quad s = \pm \frac{1}{2}. \quad (5.86)$$

Zauważmy, że blisko granicy nierelatywistycznej, kiedy prędkość $|\mathbf{v}| = \frac{|\mathbf{p}|}{E} \ll 1$, spinory χ_s są małe w tym sensie, że $\chi_s^\dagger \chi_s \ll 1$. To jest ważna w zastosowaniach własność funkcji falowych cząstek swobodnych w reprezentacji Diraca.

Funkcję $\psi_{+,s}(x)$ normalizujemy jak w poprzednim paragrafie, kładąc

$$\bar{\psi}_{+,s}\psi_{+,s} = 2. \quad (5.87)$$

Iloczyn

$$2 = \bar{\psi}_{+,s}\psi_{+,s} = N_+^2 \left(\phi_s^\dagger, \phi_s^\dagger \frac{\boldsymbol{\sigma}\mathbf{p}}{E+m} \right) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \phi_s \\ \frac{\boldsymbol{\sigma}\mathbf{p}}{E+m}\phi_s \end{pmatrix} \quad (5.88)$$

wyliczamy, uwzględniając że

$$(\boldsymbol{\sigma}\mathbf{p})^2 = \mathbf{p}^2 = E^2 - m^2 = (E+m)(E-m). \quad (5.89)$$

Rozwiązując na N_+ , otrzymujemy:

$$N_+ = \sqrt{\frac{E+m}{m}}. \quad (5.90)$$

dinopl

Możemy więc wypisać pierwsze dwa znormalizowane rozwiązania:

$$\psi_{+,s}(x) = \sqrt{\frac{E+m}{m}} \begin{pmatrix} \phi_s \\ \frac{\boldsymbol{\sigma}\mathbf{p}}{E+m}\phi_s \end{pmatrix} e^{-ipx}, \quad p_0 > 0, \quad s = \pm\frac{1}{2}. \quad (5.91)$$

Przechodzimy teraz do przypadku

$$p_0 = -E < 0. \quad (5.92)$$

defettl

Tym razem wybieramy spinory χ_s spełniające związki

$$\sigma_z \chi_s = 2s \chi_s, \quad \chi_s^\dagger \chi_s = 1. \quad (5.93)$$

Teraz s oznacza więc rzut spinu cząstki z ujemną energią na oś z . Przypominamy, że rzut spinu na stałą oś, w przeciwieństwie do skrętności, powinien zmienić znak przy sprzężeniu ładunkowym. Wyliczając spinor ϕ_s z pierwszego z równań (5.83) otrzymujemy

$$\psi_{-,s} = N_- \begin{pmatrix} -\frac{\boldsymbol{\sigma}\mathbf{p}}{E+m}\chi_s \\ \chi_s \end{pmatrix} e^{-ipx}, \quad p_0 < 0, \quad s = \pm\frac{1}{2}, \quad (5.94)$$

Rachunek bardzo podobny do poprzedniego daje

$$\bar{\psi}_{-,s}\psi_{-,s} = -N_-^2 \frac{2m}{E+m}. \quad (5.95)$$

Normalizacja do liczby dodatniej jest więc niemożliwa. Normalizując do liczby -2 , otrzymujemy

$$N_- = N_+ = \sqrt{\frac{E+m}{m}}, \quad (5.96)$$

dirneg

i w końcu

$$\psi_{-,s} = \sqrt{\frac{E+m}{m}} \begin{pmatrix} -\frac{\boldsymbol{\sigma}\mathbf{p}}{E+m}\chi_s \\ \chi_s \end{pmatrix} e^{-ipx}, \quad p_0 < 0, \quad s = \pm\frac{1}{2}. \quad (5.97)$$

Wzory (5.91) i (5.97) dają wszystkie cztery rozwiązania równania Diraca dla cząstki swobodnej w reprezentacji Diraca. Zauważmy jeszcze, że komplet czterech rozwiązań można by też otrzymać mnożąc rozwiązania otrzymane w reprezentacji spinorowej przez macierz U daną wzorem (5.78).

Ćwiczenie Powtórzyć to wyprowadzenie, wyrażając spinory χ_s przez spinory ϕ_s . Zwrócić uwagę, że wtedy pojawia się mianownik dążący do zera w granicy nierelatywistycznej, czyli kiedy reprezentacja Diraca jest szczególnie wygodna. Nie jest to błąd, bo licznik także dąży do zera i otrzymujemy wynik skończony, ale czasem komplikuje rachunki. Dlatego, w miarę możliwości opłaca się takich mianowników unikać.

•••••

Sprawdźmy, że sprzężenie ładunkowe przeprowadza rozwiązania (5.97) w rozwiązania (5.86). Na podstawie wyników z paragrafu 5.4, macierz U_C w reprezentacji Diraca dana jest wzorem

$$U_C = i\gamma_D^y = \begin{pmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{pmatrix}. \quad (5.98)$$

Stąd

$$\psi_{-,s}^C(x) = \begin{pmatrix} 0 & \epsilon \\ -\epsilon & 0 \end{pmatrix} \sqrt{\frac{E+m}{m}} \begin{pmatrix} -\frac{\boldsymbol{\sigma}^* \mathbf{p}}{E+m} \chi_s^* \\ \chi_s^* \end{pmatrix} e^{+ipx}. \quad (5.99)$$

Wykonując mnożenie macierzowe i pamiętając, że według wzoru (5.72), $\epsilon\sigma_i^* = -\sigma_i\epsilon$ otrzymujemy

$$\psi_{-,s}^C(x) = \sqrt{\frac{E+m}{m}} \begin{pmatrix} +\epsilon\chi_s^* \\ -\frac{\boldsymbol{\sigma} \mathbf{p}}{E+m} \epsilon\chi_s^* \end{pmatrix} e^{+ipx}. \quad (5.100)$$

Biorąc sprzężenie zespolone pierwszej równości we wzorze (5.93), potem mnożąc stronami lewostronnie otrzymaną równość przez macierz ϵ mamy

$$\sigma_z \epsilon \chi_s^* = -2s \epsilon \chi_s^*. \quad (5.101)$$

Porównując to równanie z pierwszym równaniem we wzorze (5.85) widać, że można położyć

$$\epsilon \chi_s^* = \phi_{-s}. \quad (5.102)$$

Robiąc w funkcji ψ^C to podstawienie i podstawienie

$$p^\mu = -p_C^\mu \quad (5.103)$$

otrzymujemy

$$\psi_s^C(x) = \sqrt{\frac{E+m}{m}} \begin{pmatrix} +\phi_{-s} \\ \frac{\boldsymbol{\sigma} \mathbf{p}_c}{E+m} \phi_{-s} \end{pmatrix} e^{-ip_C x}, \quad (5.104)$$

a zatem rozwiązanie równania dla cząstki z dodatnią energią $p_{C0} = -p_0$, z pędem $\mathbf{p}_C = -\mathbf{p}$ i z rzutem spinu na oś z równym $-s$.

5.7 Poprawki relatywistyczne do opisu cząstki w polu elektromagnetycznym: pierwsze przybliżenie

W tym paragrafie i w następnym rozpatrzemy wynikające z równania Diraca poprawki relatywistyczne do opisu cząstki o spinie $S = \frac{1}{2}$, oddziałującej z niezależnym od czasu, słabym polem elektromagnetycznym. Przyjmujemy sprzężenie minimalne. Ograniczamy się do obszaru niskich energii i do rozwiązań opisujących cząstki. Będziemy więc dyskutowali rozwiązania z $p^0 > 0$, dla równania Diraca w reprezentacji Diraca. W obecności zewnętrznego pola niezależnego od czasu, operator w równaniu nie jest już na ogół niezmienniczy względem przesunięć przestrzennych, pozostaje natomiast niezmienniczość względem przesunięć w czasie. Wobec tego funkcję falową można zapisać w postaci (3.14)

$$\psi(\mathbf{x}, t) = \begin{pmatrix} \psi_A(\mathbf{x}) \\ \psi_B(\mathbf{x}) \end{pmatrix} e^{-iEt}, \quad E = p^0 > 0. \quad (5.105)$$

Wprowadzając do równań (5.83) sprzężone minimalnie pole elektromagnetyczne, możemy napisać potrzebne nam równanie Diraca jako

$$\begin{aligned} (E - qA_0 - m)\psi_A(\mathbf{x}) - \boldsymbol{\sigma}(\hat{\mathbf{p}} - q\mathbf{A})\psi_B(\mathbf{x}) &= 0, \\ \boldsymbol{\sigma}(\hat{\mathbf{p}} - q\mathbf{A})\psi_A(\mathbf{x}) - (E - qA_0 + m)\psi_B(\mathbf{x}) &= 0, \end{aligned} \quad (5.106)$$

gdzie czterowektor A jest potencjałem wektorowym pola elektromagnetycznego, m masą cząstki, a q jej ładunkiem elektrycznym. W szczególności dla elektronu $q = -e$.

Rozważamy sytuację, kiedy efekty relatywistyczne dają tylko niewielkie poprawki do wyników nierelatywistycznych. Żeby tak było, jedynym dużym parametrem w problemie musi być masa cząstki m . Inne wielkości o wymiarze energii, $|\mathbf{p}|$ i $|qA^\mu|$, muszą być znacznie mniejsze. Energia E jest bliska m , ale energia nierelatywistyczna

$$E_{NR} = E - m \quad (5.107)$$

jest rzędu $\frac{\mathbf{p}^2}{2m}$ i co za tym idzie też znacznie mniejsza od m . Metoda, którą zastosujemy, polega na wyeliminowaniu z równań "małej" składowej $\psi_B(\mathbf{x})$ i rozwinięciu operatora w równaniu na dużą składową w szereg potęgowy w małym parametrze $\frac{1}{m}$.

Ćwiczenie Przez odpowiedni dobór jednostek można spowodować, że parametr $\frac{1}{m}$ będzie dowolnie duży. Wyjaśnić dlaczego mimo to rozważane tu rozwinięcie ma sens.

•••••

Rozwiązując drugie równanie na funkcję $\psi_B(\mathbf{x})$ i podstawiając wynik do pierwszego równania otrzymujemy

$$(E - qA_0 - m)\psi_A(\mathbf{x}) - \boldsymbol{\sigma}(\hat{\mathbf{p}} - q\mathbf{A}) \frac{1}{E - qA_0 + m} \boldsymbol{\sigma}(\hat{\mathbf{p}} - q\mathbf{A})\psi_A(\mathbf{x}) = 0. \quad (5.108)$$

Zauważmy, że funkcja $(E - qA_0 + m)^{-1}$ na ogół nie komutuje z operatorem $\sigma(\hat{\mathbf{p}} - q\mathbf{A})$, więc kolejność czynników jest tu istotna. To równanie można przepisać w postaci podobnej do równania Schrödingera niezależnego od czasu

$$\hat{H}\psi_A(x) = E_{NR}\psi_A(x), \quad \hat{H} = \sigma(\hat{\mathbf{p}} - q\mathbf{A}) \frac{1}{E_{NR} - qA_0 + 2m} \sigma(\hat{\mathbf{p}} - q\mathbf{A}) + qA^0. \quad (5.109)$$

Drugi człon w hamiltonianie daje po prostu energię potencjalną cząstki punktowej o ładunku q w polu potencjału elektrostatycznego A^0 . Pierwszy jest trudniejszy do interpretacji.

Pierwsze przybliżenie otrzymujemy kładąc

$$\frac{1}{E_{NR} - qA_0 + 2m} = \frac{1}{2m} + O(m^{-2}) \quad (5.110)$$

i zaniehbując poprawkę. Wtedy

$$\hat{H} = \frac{[\sigma(\hat{\mathbf{p}} - q\mathbf{A})]^2}{2m} + qA^0. \quad (5.111)$$

Ten hamiltonian różni się od hamiltonianu wprowadzonego w rozdziale drugim tylko obecnością macierzy σ_j . Gdyby składowe wektora $\hat{\mathbf{a}} = \hat{\mathbf{p}} - q\mathbf{A}$ komutowały między sobą, to zachodziłaby równość $(\sigma\hat{\mathbf{a}})^2 = \hat{\mathbf{a}}^2$ i wrócilibyśmy do tamtego równania. Ale w ogólnym przypadku, jak łatwo sprawdzić,

$$(\sigma\hat{\mathbf{a}})^2 = \hat{\mathbf{a}}^2 + i\sigma(\hat{\mathbf{a}} \times \hat{\mathbf{a}}). \quad (5.112)$$

W rozważanym przybliżeniu możemy napisać hamiltonian w postaci

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \quad (5.113)$$

gdzie \hat{H}_0 jest hamiltonianem z rozdziału drugiego, a \hat{H}_1 poprawką pochodzącą od członu z iloczynem wektorowym operatorów $\hat{\mathbf{a}}$. Pozostaje do obliczenia

$$[(\hat{\mathbf{p}} - q\mathbf{A}) \times (\hat{\mathbf{p}} - q\mathbf{A})] = iq(\nabla \times \mathbf{A} + \mathbf{A} \times \nabla), \quad (5.114)$$

gdzie uwzględniliśmy fakt, że iloczyn wektorowy dwu identycznych wektorów znika, jeśli składowe wektora komutują między sobą, i wstawiliśmy wyrażenia na operatory składowych pędu. Operatory różniczkowe w pierwszym wyrazie po prawej stronie działają nie tylko na wektor \mathbf{A} , ale i na wszystko, co się znajduje na prawo od niego. Możemy więc napisać

$$\nabla \times \mathbf{A} = \mathbf{B} - \mathbf{A} \times \nabla, \quad (5.115)$$

gdzie wykorzystaliśmy tożsamość $\mathbf{B} = (\nabla \times \mathbf{A})$ i antysymetrię iloczynu wektorowego. Pamiętając, że spin elektronu $\mathbf{S} = \frac{1}{2}\sigma$ otrzymujemy

$$\hat{H}_1 = -\frac{q}{m}\mathbf{S}\mathbf{B}. \quad (5.116)$$

Ta poprawka do hamiltonianu jest znana jako człon Pauli'ego i była już dyskutowana w paragrafie 2.4.

Dla zwykle spotykanych pól poprawka do energii wynikająca z członu \hat{H}_1 jest niewielka. Na przykład, dla elektronu w polu o natężeniu jednej tesli (10^4

gaussów) wynosi około $\pm 5.8 \times 10^{-5}$ eV, do porównania z różnicą energii między stanem podstawowym i pierwszym stanem wzbudzonym atomu wodoru, która wynosi około 10 eV. Wyjątkiem od tej praktycznej reguły są atomy na powierzchni gwiazdy neutronowej, ale tam pole magnetyczne jest tak silne, że podana tu prosta metoda uwzględnienia tego pola zawodzi. Rozszczepienie poziomów energetycznych atomów pod wpływem pola magnetycznego jest znane jako efekt Zeemana, ale w praktyce zwykle ma się do czynienia z tak zwanym anomalnym efektem Zeemana, kiedy rozszczepienia pochodzące od poprawki (5.116) nakładają się z innymi, porównywalnymi co do wielkości, poprawkami, które dyskutujemy w następnym paragrafie.

Porównując człon Pauli'ego z ogólnym wzorem na energię cząstki z momentem magnetycznym $\boldsymbol{\mu}$ w polu magnetycznym o natężeniu \mathbf{B} ,

$$E_{magn} = -\boldsymbol{\mu}\mathbf{B}, \quad (5.117)$$

otrzymujemy wzór na moment magnetyczny cząstki

$$\boldsymbol{\mu} = \frac{\hbar q}{2mc} g \mathbf{S}, \quad g = 2. \quad (5.118)$$

Ten wynik był już zapowiedziany w paragrafie 2.4. Otrzymanie (prawie) poprawnej wartości współczynnika żyromagnetycznego g dla elektronu, zakładając tylko sprzężenie minimalne z polem elektromagnetycznym, było wielkim sukcesem teorii Diraca.

Pokażemy teraz, że w teoriach, w których obowiązuje zachowanie parzystości, cząstka punktowa, chociaż może mieć moment (dipolowy) magnetyczny, nie może mieć momentu dipolowego elektrycznego. W takich teoriach energia nie może zmieniać znaku przy inwersji przestrzennej. Dla energii magnetycznej pseudowektor \mathbf{B} jest mnożony przez moment magnetyczny proporcjonalny do pseudowektora \mathbf{S} , więc ten warunek jest spełniony. Pole elektryczne \mathbf{E} jest zwykłym wektorem. Ewentualny moment dipolowy elektryczny musiałby więc też być zwykłym wektorem. Ale dla cząstki punktowej w spoczynku nie ma żadnego różnego od zera wektora opisującego jej stan ($\mathbf{p} = \mathbf{0}$). Stąd nie może też być różnego od zera momentu dipolowego elektrycznego. Jeżeli teoria nie zachowuje parzystości, to nie ma podstaw do twierdzenia, że energia nie zmienia się przy inwersji przestrzennej. Wtedy może, choć nie musi, wystąpić różny od zera elektryczny moment dipolowy. Teoria słabych oddziaływań nie zachowuje parzystości. Ze względu na to trwają poszukiwania dipolowego momentu elektrycznego neutronu. Na razie otrzymuje się tylko górne granice. Obecnie wiadomo [15], że ten moment jest mniejszy niż około 3×10^{-26} e cm.

5.8 Poprawki relatywistyczne do opisu cząstki w polu elektromagnetycznym: drugie przybliżenie

W drugim przybliżeniu rachunek, którego tu nie pokazujemy³, daje, przy założeniu $\mathbf{B} = \mathbf{0}$, dodatkowy przyczynek do hamiltonianu

³Ten rachunek można znaleźć na przykład w podręczniku [16]

$$\hat{H}_2 = \hat{H}_{2,1} + \hat{H}_{2,2} + \hat{H}_{2,3}, \quad (5.119)$$

gdzie

$$\hat{H}_{2,1} = -\frac{(\hat{\mathbf{p}}^2)^2}{8m^3}, \quad (5.120)$$

$$\hat{H}_{2,2} = -\frac{q}{4m^2}\boldsymbol{\sigma}(\mathbf{E} \times \hat{\mathbf{p}}), \quad (5.121)$$

$$\hat{H}_{2,3} = -\frac{q}{8m^2}(\nabla \cdot \mathbf{E}). \quad (5.122)$$

W tych wzorach \mathbf{E} jest natężeniem pola elektrycznego. Nawias w ostatnim wzorze zaznacza, że operator nabla działa tylko na wektor \mathbf{E} , a nie na dalsze wyrażenia, które mogłyby być dopisane dalej na prawo. Ładunek elektronu $q = -e$ wiąże się ze stałą struktury subtelnej

$$\alpha = e^2 \approx \frac{1}{137}. \quad (5.123)$$

Ta stała jest bezwymiarowa. Wszyscy się zgadzają co do jej wartości numerycznej, ale związek z ładunkiem elektronu zależy od przyjętego układu jednostek. Tu, jak zwykle w mechanice kwantowej, używamy systemu Gaussa, ale spotyka się także $\alpha = \frac{e^2}{4\pi}$ (układ Heaviside'a-Lorentza), czy $\alpha = 10^{-7}e^2$ (układ MKSA). Bardzo dobry, bezstronny opis różnych układów jednostek używanych w elektrodynamice można znaleźć w podręczniku Jacksona [9]. Poprawka do energii układu opisywanego przez funkcję falową $\psi(x)$, pochodząca od poprawki do hamiltonianu \hat{H}_2 , jest w bardzo dobrym przybliżeniu dana wziętym z metody zaburzeń wzorem

mepert

$$\Delta E_2 = \int dx \psi^*(x) \hat{H}_2 \psi(x), \quad (5.124)$$

gdzie funkcje falowe $\psi(x)$ są rozwiązaniami nierelatywistycznego równania Schrödingera. Przedyskutujemy teraz interpretację fizyczną trzech poprawek $\hat{H}_{2,i}$.

Rozwijając relatywistyczny wzór na energię kinetyczną w szereg potęg małego parametru $\frac{\mathbf{p}^2}{m^2}$ otrzymujemy

$$E(\mathbf{p}) = \sqrt{\mathbf{p}^2 + m^2} = m + \frac{\mathbf{p}^2}{2m} - \frac{(\mathbf{p}^2)^2}{8m^3} + \dots \quad (5.125)$$

Przyczynek $\hat{H}_{2,1}$ jest więc relatywistyczną **poprawką do nierelatywistycznego wzoru na energię kinetyczną** cząstki. W stanie podstawowym i niskich stanach wzbudzonych atomu wodoru, prędkość elektronu $\frac{|\mathbf{p}|}{m}$ jest rzędu α , stąd poprawka $\hat{H}_{2,1}$ jest rzędu $m\alpha^4$.

Poprawkę $\hat{H}_{2,2}$ przedyskutujemy dla sferycznie symetrycznego potencjału elektrostatycznego, gdzie

$$\mathbf{E} = -\nabla V(r) = -\frac{\mathbf{x}}{r} \frac{\partial V(r)}{\partial r}. \quad (5.126)$$

Podstawiając to wyrażenie do wzoru na $\hat{H}_{2,2}$ otrzymujemy

$$\hat{H}_{2,2} = \frac{q}{4m^2r} \frac{\partial V(r)}{\partial r} \boldsymbol{\sigma}(\mathbf{x} \times \hat{\mathbf{p}}) = \frac{q}{2m^2r} \frac{\partial V(r)}{\partial r} \mathbf{S}\hat{\mathbf{L}}. \quad (5.127)$$

To jest **sprzężenie spin-orbita**, odpowiedzialne za strukturę subtelną linii widmowych. Nierelatywistyczne funkcje falowe we wzorze (5.124) wybieramy tak, żeby były wspólnymi funkcjami własnymi operatorów \mathbf{J}^2 , \mathbf{L}^2 i \mathbf{S}^2 . Ponieważ $\mathbf{J} = \mathbf{L} + \mathbf{S}$, operator $\mathbf{L}\mathbf{S}$ też jest diagonalny w tej reprezentacji i można go zapisać w postaci:

$$\mathbf{L}\mathbf{S} = \frac{1}{2} \left[J(J+1) - L(L+1) - \frac{3}{4} \right] = \frac{1}{2} [(J-L)(J+L+1) - \frac{3}{4}], \quad (5.128)$$

gdzie uwzględniliśmy fakt, że $\mathbf{S}^2 = S(S+1) = \frac{3}{4}$. Sprzężenie spin-orbita wprowadza zależność energii stanu od tego, czy w tym stanie spin jest równoległy ($J = L + \frac{1}{2}$), czy też antyrównoległy ($J = L - \frac{1}{2}$) do krętu orbitalnego L . Na skutek tego, każdy poziom schrödingerski, za wyjątkiem poziomów z $L = 0$ dla których sprzężenie spin-orbita zeruje się, zostaje rozszczepiony na dwa bliskie sobie poziomy. W szczególności, podstawiając za potencjał $qV(r)$ potencjał coulombowski $-\frac{\alpha}{r}$ otrzymujemy

$$\hat{H}_{2,2} = \frac{\alpha}{2m^2r^3} \mathbf{S}\mathbf{L} \quad (5.129)$$

Przyjmując, że typowy promień r jest rzędu odwrotności typowego pędu $|\mathbf{p}|$, którą podaliśmy przy dyskusji poprawki $H_{2,1}$, mamy $r \approx \frac{1}{m\alpha}$. Widać, że i ta poprawka jest rzędu $m\alpha^4$. Sprzężenie spin-orbita ma prostą interpretację fizyczną. Elektron krążąc wytwarza, jak każdy krążący ładunek elektryczny, pole magnetyczne. To pole oddziałuje z momentem magnetycznym elektronu.

Przyczynę $\hat{H}_{2,3}$ nazywa się **członem Darwina**. Zauważmy, że według równań Maxwella dywergencja z natężenia pola elektrycznego, $\nabla \cdot \mathbf{E}$, jest równa pomnożonej przez 4π gęstości ładunku wytwarzającego pole. Zakładając, że źródłem pola jest punktowy proton spoczywający w początku układu współrzędnych, mamy więc

$$\nabla \cdot \mathbf{E} = 4\pi|e|\delta^3(\mathbf{x}). \quad (5.130)$$

Podstawiając do wzoru na $\hat{H}_{2,3}$, otrzymujemy

$$\hat{H}_{2,3} = \frac{\pi\alpha}{2m^2} \delta^3(\mathbf{x}). \quad (5.131)$$

Temu wyrażeniu odpowiada zmiana energii o

$$\int dx \psi^*(\mathbf{x}) \hat{H}_{2,3} \psi(\mathbf{x}) = \frac{\pi\alpha}{2m^2} |\psi(0)|^2, \quad (5.132)$$

gdzie całkowanie wykonaliśmy wykorzystując deltę Diraca. Dla wszystkich funkcji falowych stanów stacjonarnych atomu wodoru z $L \neq 0$: $\psi(\mathbf{0}) = \mathbf{0}$. Poprawka Darwina jest więc różna od zera tylko dla tak zwanych stanów S , czyli dla stanów z $L = 0$. Dla gęstości $|\psi(\mathbf{x})|^2$ rozłożonej równomiernie wewnątrz kostki, której długość krawędzi jest równa typowemu r , czyli $\frac{1}{m\alpha}$, mielibyśmy gęstość w środku

$$|\psi(0)|^2 \approx m^3 \alpha^3. \quad (5.133)$$

przyjmując to oszacowanie także dla stanu podstawowego i niewysoko wzbudzonych stanów S atomu wodoru widzimy, że człon Darwina, podobnie jak dwa poprzednie, daje przyczynek do energii rzędu $m\alpha^4$.

Człon Darwina można rozumieć jako wynik rozmycia elektronu w chmurę o promieniu rzędu $\frac{1}{m}$, które omawialiśmy w paragrafie 5.3 (twierdzenie Newtona i Wignera). To oszacowanie promienia można by też oprzeć na zasadzie nieokreśloności dla energii i czasu. Fluktuacja polegająca na powstaniu pary e^+e^- zwiększa energię o około $2m$, a zatem trwa przez czas rzędu $\frac{1}{m}$. Przez ten czas cząstki mogą oddalić się od siebie w układzie swojego środka masy na odległość rzędu $\frac{1}{m}$. Pokażemy teraz, jak proste oszacowanie oparte na założeniu takiego rozmycia pozwala w przybliżeniu odtworzyć $\hat{H}_{2,3}$. To oszacowanie może się wydawać naiwne, ale należy pamiętać, że chodzi nam tylko o rząd wielkości współczynnika i ogólną postać wzoru.

Przyjmijmy dla chmury, w którą rozmyty jest elektron,

$$\langle \delta x_i \rangle = 0, \quad \langle \delta x_i \delta x_j \rangle = \frac{\delta_{ij}}{3m^2}, \quad (5.134)$$

gdzie δ_{ij} jest deltą Kroneckera, przesunięcia $\delta \mathbf{x}$ są liczone względem środka chmury a średniowania są po chmurze. Trójka w mianowniku jest potrzebna, bo $(\delta \mathbf{x})^2 = (\delta x_1)^2 + (\delta x_2)^2 + (\delta x_3)^2$. Obliczymy teraz energię potencjalną oddziaływania tej chmury z protonem. Ponieważ chmura jest mała, ograniczymy się do przybliżenia z dokładnością do członów drugiego rzędu w składowych $\delta \mathbf{x}$:

$$\langle V(\mathbf{x} + \delta \mathbf{x}) \rangle = V(\mathbf{x}) + \sum_{i=1}^3 \frac{\partial V(\mathbf{x})}{\partial x_i} \langle \delta x_i \rangle + \frac{1}{2} \sum_{j,k} \frac{\partial^2 V(\mathbf{x})}{\partial x_j \partial x_k} \langle \delta x_j \delta x_k \rangle, \quad (5.135)$$

gdzie wszystkie pochodne są liczone w punkcie \mathbf{x} . Stąd

$$\langle V(\mathbf{x} + \delta \mathbf{x}) \rangle - V(\mathbf{x}) = \frac{1}{6m^2} \nabla^2 V(\mathbf{x}) = -\frac{q}{6m^2} \nabla \cdot \mathbf{E}. \quad (5.136)$$

Poprawka do energii potencjalnej ma strukturę członu Darwina i stały współczynnik też został nieźle odtworzony.

Na koniec przedyskutujemy poprawki, relatywistyczne, i niektóre inne, do schrödingerowskich poziomów energetycznych elektronu w atomie wodoru. Dodając do wartości własnych energii otrzymanych z nierelatywistycznego równania Schrödingera wartości średnie poprawek do hamiltonianu $\hat{H}_{2,i}$, $i = 1, 2, 3$, otrzymujemy

dirhyd

$$E_{NR}(n, J) = -\frac{m\alpha^2}{2n^2} + \frac{m\alpha^4}{n^4} \left(\frac{3}{8} - \frac{n}{2J+1} \right), \quad (5.137)$$

gdzie $n = 1, 2, \dots$ jest tak zwaną główną liczbą kwantową, a $J = L \pm \frac{1}{2}$ oznacza całkowity (orbitalny + spin) kręt elektronu. Ten sam wzór można uzyskać rozwiązując dokładnie równanie Diraca dla elektronu w polu kulombowskim jądra, rozwijając wzór na wartości własne energii w szereg potęgowy w α i zaniedbując człony rzędów wyższych niż α^4 . Z dokładnego rozwiązania można uzyskać dalsze człony tego rozwinięcia, ale porównanie z wynikami elektrodynamiki kwantowej, lub z doświadczeniem, wskazuje, że te dalsze człony nie są już poprawne, więc ograniczymy się do pierwszych dwu. Podstawiając za masę elektronu 0,511 MeV otrzymujemy

$$m\alpha^2 \approx 27,2\text{eV}, \quad m\alpha^4 \approx 0,00145\text{eV}, \quad (5.138)$$

więc poprawki są nieduże. Mimo to są ważne, bo rozszczepienia poziomów zmieniają jakościowo widmo.

Dla stanów S ($L = 0$) sprzężenie spin-orbita nie daje przyczynku i w rozważanym przybliżeniu poziomy się nie rozszczepiają, choć nieco się przesuwają dzięki poprawce do energii kinetycznej i członowi Darwina. Dla stanów z $L \neq 0$ są dwie możliwe wartości J i odpowiadają im dwie różne energie — poziomy zostają rozszczepione. Te rozszczepienia są nazywane strukturą subtelną poziomów energetycznych, i co za tym idzie linii widmowych emitowanego lub pochłanianego światła. Na przykład dla stanu $2P$, ($n = 2$, $L = 1$) ze wzoru (5.137) otrzymujemy wielkość tego rozszczepienia $\Delta E \approx 4,53 \times 10^{-5}\text{eV}$, co odpowiada długości fali $\lambda \approx 2,74\text{cm}$, a więc w zakresie mikrofal. Te wyniki zupełnie dobrze zgadzają się z doświadczeniem, ale są dwie bardzo ważne, choć numerycznie nieduże, rozbieżności. Poziomy S mają jednak niewielkie rozszczepienia. W szczególności dla stanu $1S_{\frac{1}{2}}$ rozszczepienie wynosi około $5.9 \times 10^{-6}\text{eV}$, co odpowiada fali o długości 21cm . To jest tak zwane **rozszczepienie nadsubtelne**, widoczne i dla innych poziomów, które wynika z oddziaływania momentu magnetycznego elektronu z momentem magnetycznym jądra, w tym przypadku protonu. Linia 21cm w promieniowaniu elektromagnetycznym dochodzącym z przestrzeni kosmicznej do Ziemi ma wielkie znaczenie dla astronomii, bo jest słabo absorbowana w przestrzeni międzygwiazdnej i jej natężenie pozwala oceniać ilości wodoru w odległych obszarach kosmosu, a jej döpplerowskie przesunięcie określa prędkość emitującej chmury wodoru. Historia jej zaobserwowania – jednego z wielkich sukcesów uczonych, którzy w czasie wojny pracowali nad radarem – jest interesująco opisana w książce [17]. Ponieważ przy danym n energia zależy tylko od J , równanie Diraca przewiduje, że stany $2S_{\frac{1}{2}}$ i $2P_{\frac{1}{2}}$ powinny mieć tę samą energię. Doświadczalnie obserwuje się tak zwane **przesunięcie Lamba** – energia stanu $2P_{\frac{1}{2}}$ jest niższa o około $4 \times 10^{-6}\text{eV}$, co odpowiada długości fali około 29cm , lub częstotliwości około 1040MHz . Wyjaśnienie tego efektu, podobnie jak wyliczenie anomalnego momentu magnetycznego elektronu, wymaga zastosowania elektrodynamiki kwantowej, a więc przekracza możliwości równania Diraca. Zauważmy, że przesunięcie Lamba jest tylko o rząd wielkości mniejsze od rozszczepienia subtelnego, więc uwzględnianie znacznie mniejszego następnego członu w rozwinięciu (5.137) byłoby nieuzasadnione. Kwantowa teoria pola pozwala na zgodne z doświadczeniem wyliczenie przesunięcia Lamba.

5.9 Paradoks Kleina

Jako kolejny przykład zastosowania równania Diraca, rozważymy problem jednowymiarowy, z minimalnym sprzężeniem elektronu z polem elektromagnetycznym, które ma tylko składową A^0 potencjału wektorowego różną od zera. Dla tej składowej przyjmijmy: $qA^0 = V(z)$, gdzie $V(z) = 0$ dla $z < 0$ i $V(z) = V_0 > 2m$ dla $z > 0$. Gradient potencjału, czyli siła, jest więc nieskończony w punkcie $z = 0$ i znika poza tym punktem. Jest to oczywiście daleko idąca idealizacja, ale prowadzi do ciekawych wniosków. Rozwiązując problem rozpraszania elektronu na progu potencjału w $z = 0$ dojdziemy do tak zwanego paradoksu Kleina.

Szczególnie ciekawe wyniki otrzymuje się, jeśli energia cząstki E spełnia nierówność

$$m < E < V_0 - m. \quad (5.139)$$

Dalej rozpatrujemy tylko ten przypadek. Elektron nadlatujący z pędem $p_z > 0$ i rzutem spinu na oś z równym $+\frac{1}{2}$ może odbić się, ze zmianą kierunku spinu lub bez, albo może polecieć dalej, znów ze zmianą kierunku spinu lub bez. Klasycznie, przejście elektronu do obszaru $z > 0$ jest oczywiście zakazane przez zasadę zachowania energii. Problem rozpraszania rozwiążemy metodą stanów stacjonarnych⁴, czyli znajdując rozwiązanie niezależnego od czasu równania Schrödingera i odpowiednio je interpretując. Użyjemy reprezentacji Diraca. Rozwiązania uzyskane w paragrafie 5.6 odpowiadają stałej energii, a więc stanom stacjonarnym. W tym paragrafie będziemy je stosować pomijając czynniki e^{-ip_0t} , które się wszędzie upraszczają.

Warunki brzegowe muszą odpowiadać rozważanej sytuacji fizycznej. W naszym przypadku, dla $z \rightarrow -\infty$ chcemy mieć funkcję falową cząstki padającej $\psi_{in+}(z)$, czyli cząstki swobodnej z pędem $p_z = +\sqrt{E^2 - m^2}$ i $S_z = +\frac{1}{2}$ oraz funkcje falowe cząstek odbitych $\psi_{R\pm}(x)$ z pędem $-p_z$ i rzutem spinu na oś z równym odpowiednio $+\frac{1}{2}$ lub $-\frac{1}{2}$. Ponieważ cząstki pozostają swobodne w całym obszarze $z < 0$, mamy dla całego tego przedziału funkcję falową

$$\psi(z) = u_{in+}(p_z)e^{ip_z z} + R_+ u_{R+}(-p_z)e^{-ip_z z} + R_- u_{R-}(-p_z)e^{-ip_z z}, \quad z < 0, \quad (5.140)$$

gdzie stałe R_{\pm} wyznacza się z równania. Zauważmy, że ponieważ równanie Diraca jest jednorodne, możemy wybrać współczynnik przy u_{in+} równy jeden i to dla dowolnej normalizacji tej funkcji. Normalizacje bispinorów $u_{R\pm}$ są też dowolne, bo ich zmiany można skompensować odpowiednimi zmianami współczynników R_{\pm} . Uwzględniając to wszystko i korzystając z wyników paragrafu 5.6 dla reprezentacji Diraca otrzymujemy dla obszaru $z < 0$

$$\psi(z) = \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ 0 \end{pmatrix} e^{ip_z z} + R_+ \begin{pmatrix} 1 \\ 0 \\ \frac{-p_z}{E+m} \\ 0 \end{pmatrix} e^{-ip_z z} + R_- \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{p_z}{E+m} \end{pmatrix} e^{-ip_z z}, \quad z < 0. \quad (5.141)$$

W obszarze $z > 0$ mamy cząstki, które przeleciały przez próg potencjału i lecą w kierunku $+z$, bez zmiany energii E i z rzutami spinu na oś z równymi odpowiednio $\pm\frac{1}{2}$. Ich funkcje falowe są funkcjami falowymi cząstek poruszających się w polu stałego potencjału V_0 , będącego zerową składową czterowektora. Te funkcje otrzymuje się z funkcji falowych cząstek swobodnych zastępując wszędzie E przez $E - V_0$. Pęd p_z zostaje więc zastąpiony przez

$$p'_z = \sqrt{(E - V_0)^2 - m^2}. \quad (5.142)$$

Zauważmy, że w interesującym nas obszarze energii funkcja $p'_z(E, V_0, m)$ jest rzeczywista. Znak pierwiastka wynika z przyjętych warunków brzegowych. Funkcja falowa ma postać

⁴Ta metoda jest dla naszego przypadku poprawna, ale jej uzasadnienie, które kiedyś w podręcznikach przedstawiano jako oczywiste, jest w rzeczywistości dość skomplikowane, nawet dla przypadku rozpraszania nierelatywistycznego. Dobrą dyskusję tego zagadnienia można znaleźć w monografii [12].

$$\psi(z) = T_+ \begin{pmatrix} 1 \\ 0 \\ \frac{p'_z}{E-V_0+m} \\ 0 \end{pmatrix} e^{ip'_z z} + T_- \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{-p'_z}{E-V_0+m} \end{pmatrix} e^{ip'_z z}, \quad z > 0, \quad (5.143)$$

gdzie znów stałe T_{\pm} należy wyznaczyć z równania.

Żeby otrzymać funkcję falową ważną w całym obszarze $-\infty < z < \infty$, trzeba dobrać stałe R_{\pm} i T_{\pm} tak, żeby funkcja $\psi(z)$ była ciągła w punkcie $z = 0$. Wygodnie jest wprowadzić parametr

$$r = \frac{p'_z}{p_z} \frac{E+m}{E-V_0+m}. \quad (5.144)$$

Zauważmy, że w interesującym nas obszarze energii, r jest ujemne. Podstawiając wzory na p_z i p'_z oraz upraszczając ułamki otrzymujemy

$$r = -\sqrt{\frac{(E+m)(V_0-E+m)}{(E-m)(V_0-E-m)}}. \quad (5.145)$$

Ponieważ pod pierwiastkiem licznik $(m^2 - E^2 + V_0E + V_0m)$ jest większy od mianownika $(m^2 - E^2 + V_0E - V_0m)$, mamy nierówność

$$r < -1. \quad (5.146)$$

Warunki ciągłości dla czterech składowych bispinora w punkcie $z = 0$ dają cztery równania:

$$\begin{aligned} 1 + R_+ &= T_+ \\ R_- &= T_- \\ 1 - R_+ &= rT_+ \\ R_- &= -rT_-, \end{aligned} \quad (5.147)$$

gdzie równania trzecie i czwarte zostały podzielone stronami przez $\frac{p_z}{E+M}$. Wobec $r < -1$, jedynym rozwiązaniem równań drugiego i czwartego jest

$$R_- = T_- = 0. \quad (5.148)$$

Można było przewidzieć, że orientacja spinu nie może się zmienić, bo hamiltonian nie zależy od spinu. Rozwiązując układ pozostałych dwu równań otrzymujemy

$$T_+ = \frac{2}{1+r}, \quad R_+ = \frac{1-r}{1+r}. \quad (5.149)$$

To już jest pełne rozwiązanie. Pozostaje je zinterpretować.

Korzystając z wyników paragrafu 5.5 obliczamy z -owe składowe prądów: cząstek padających (J_{in}), cząstek odbitych (J_R) i cząstek przepuszczonych (J_T). Otrzymujemy

$$J_{in}^z = \frac{2p_z}{E+m}, \quad J_R^z = -R_+^2 \frac{2p_z}{E+m}, \quad J_T^z = T_+^2 \frac{2p'_z}{E-V_0+m}. \quad (5.150)$$

Ponieważ te strumienie są liczone z dokładnością do wspólnego stałego czynnika, ciekawsze są ich stosunki

$$\frac{J_R^z}{J_{in}^z} = -\left(\frac{1-r}{1+r}\right)^2, \quad \frac{J_T^z}{J_{in}^z} = \frac{4r}{(1+r)^2}, \quad \frac{-J_R^z + J_T^z}{J_{in}^z} = 1. \quad (5.151)$$

W trzecim wzorze zmieniliśmy znak J_R^z , żeby otrzymać całkowity strumień cząstek rozproszonych, czyli oddalających się od progu. Ten ostatni wzór można było przewidzieć: łączny strumień cząstek rozproszonych, w przód i w tył, jest równy strumieniowi cząstek padających. Pozostałe dwa wyniki są tak zaskakujące, że cały ten problem jest znany jako paradoks Kleina. Ponieważ $r < 0$, strumień cząstek rozproszonych do tyłu jest większy niż strumień cząstek padających, a strumień cząstek przepuszczonych jest ujemny! Tłumaczy się ten wynik jako przejaw produkcji par e^+e^- na skoku potencjału. Elektrony z tych par zasilają strumień odbity, a dodatni strumień pozytonów jest przedstawiony jako ujemny strumień elektronów.

Rachunek oparty na kwantowej teorii pola [14], na razie niezbyt dokładny, sugeruje, że rzeczywiście ma miejsce produkcja par na progu, ale do tego nie potrzeba padających cząstek. Wystarczy samo pole, jak w efekcie Hawkinga, który omówimy w następnym paragrafie. Padające cząstki nawet nieco tłumią tę produkcję. Równanie Diraca słusznie więc sugeruje, że na skoku potencjału jest produkcja cząstek, ale nie nadaje się do ilościowego opisu tego procesu.

5.10 Promieniowanie Hawkinga

Innym przykładem produkcji par w silnym polu, tym razem grawitacyjnym, jest promieniowanie Hawkinga. Według elektrodynamiki kwantowej próżnia jest wypełniona wirtualnymi (to znaczy takimi, że dla nich $p^2 \neq 0$) fotonami. Teoria Diraca sugeruje, że te fotony mogłyby produkować pary cząstka-antycząstka. Żeby taka wirtualna para mogła się utrwalić, trzeba jej dostarczyć energii zanim zanihiluje. Znanym przykładem jest produkcja par e^+e^- pod działaniem pola fali elektromagnetycznej, która w tym wypadku dostarcza potrzebną energię. Jeżeli energia nie zostaje dostarczona, to na mocy zasady nieokreśloności dla energii i czasu, para powstała z próżni może żyć przez czas rzędu $\frac{1}{2m}$, gdzie m jest masą cząstki, a potem musi zanihilować. Hawking zauważył, że w obszarze tuż przy horyzoncie czarnej dziury pole grawitacyjne jest dostatecznie silne, żeby dostarczyć potrzebnej energii i to w dostatecznie krótkim czasie. Wtedy jedna cząstka z pary zostaje wchłonięta przez czarną dziurę, a druga odlatuje jako cząstka której czteropęd spełnia zwykłą relację $p^2 = m^2$. Powstały w ten sposób strumień cząstek emitowanych z najbliższego otoczenia czarnej dziury nazywa się promieniowaniem Hawkinga. Dowodzi się, że emitowane cząstki mają termiczny rozkład energii, odpowiadający temperaturze znanej jako temperatura Hawkinga,

$$T_H = \frac{\hbar c^3}{4k_B M G}, \quad (5.152)$$

gdzie k_B jest stałą Boltzmana, M masą czarnej dziury, a G stałą Newtona. Podstawiając liczby, otrzymujemy w kelwinach

$$T_H \approx 3.9 \times 10^{-7} \frac{M}{M_\odot} K, \quad (5.153)$$

gdzie M_\odot oznacza masę słońca, to znaczy około 2×10^{30} kg. Jak widać, dla czarnych dziur, które nie są znacznie lżejsze od Słońca, temperatura Hawkinga jest znikomo mała.

Czarna dziura traci energię na promieniowanie Hawkinga i staje się coraz lżejsza. Kiedy jej masa maleje, temperatura Hawkinga się podnosi i energia jest wynoszona coraz prędkiej. Czarna dziura o masie około 10^6 kg ma przed sobą jeszcze sekundę życia i w tym czasie wypromieniowuje całą swoją energię, czyli około 10^{23} dżuli. W skali astronomicznej nie jest to bardzo dużo, bo na przykład Słońce w ciągu jednej sekundy wypromieniowuje energię około 4×10^{26} dżuli, ale jest to jednak potężny wybuch. Wynika stąd, że lekkie czarne dziury nie mogą żyć długo. Żeby czarna dziura miała szansę przeżyć od Wielkiego Wybuchu do naszych czasów, musi mieć masę co najmniej 10^{12} kg. Znowu nie jest to dużo w skali astronomicznej, ale wynik jest ważny dla kosmologii. Ścisłe wyprowadzenie teorii promieniowania Hawkinga musiałyby się opierać na kwantowej teorii grawitacji, a ta teoria nie została jeszcze odkryta. Używane uzasadnienia własności promieniowania Hawkinga są więc z konieczności nie całkiem ścisłe, ale uchodzą za wiarygodne.

5.11 Cząstki Majorany

Dyskutując symetrie równania Kleina-Gordona zwróciliśmy uwagę, że niektóre bozony są identyczne ze swoimi antycząstkami. Mogłoby tak być i dla fermionów (Majorana 1937). Fermiony identyczne ze swoimi antycząstkami nazywamy dziś cząstkami Majorany. Cząstką Majorany nie może być cząstka, która ma różny od zera ładunek elektryczny, bo sprzężenie ładunkowe zmienia jego znak, cząstki z różnymi ładunkami nie mogą być identyczne.

Póki wierzono w zachowanie liczby barionowej i liczby leptonowej, rozumowanie analogiczne do rozumowania dla cząstek z niezerowym ładunkiem elektrycznym prowadziło do wniosku, że fermiony nie mogą być cząstkami Majorany. Dziś powątpiewa się w prawa zachowania, które nie wynikają z jakichś zasad symetrii, a dla liczby leptonowej czy barionowej takie zasady nie są znane. Jeśli liczba barionowa czy leptonowa nie jest zachowywana, to może tak być, że jakiś barion czy lepton jest identyczny ze swoją antycząstką. Pary barion-antybarion, zderzając się przy niskiej energii, mają duże prawdopodobieństwo anihilacji. Ponieważ nie obserwuje się anihilacji par barion-barion, bariony są różne od swoich antycząstek, a więc nie są cząstkami Majorany. Natomiast hipoteza, że cząstkami Majorany są neutrino, ma wielu zwolenników, choć nie ma jeszcze żadnego dowodu doświadczalnego, że tak jest. Poza fizyką cząstek, teoria fermionów Majorany, odpowiednio zmodyfikowana, znajduje zastosowanie w teorii nadprzewodników.

Założenie, że neutrino są cząstkami Majorany ma ciekawe konsekwencje. Do najśłynniejszych należy możliwość bezneutrinowego rozpadu β . Rozważmy jądro o ładunku Z , takie że jądro o tej samej liczbie masowej i ładunku $Z + 1$ jest od niego cięższe, a jądro o tej samej liczbie masowej i ładunku $Z + 2$ lżejsze. Takie jądra istnieją. Przykładem jest ${}^{76}\text{Ge}$. W zwykłym, pojedynczym

rozpadzie β , neutron rozpada się na proton, elektron i antyneutrino, liczba masowa jądra nie zmienia się, a jego ładunek wzrasta o jeden. Z jądra wylatują elektron i antyneutrino. Dla jąder, o których tu mówimy, pojedynczy rozpad β jest zakazany przez zasadę zachowania energii. Możliwy jest natomiast, przewidziany przez panią Göppert Mayer (1935) i stwierdzony doświadczalnie, podwójny rozpad β , gdzie dwa neutrony rozpadają się, każdy na proton, elektron i antyneutrino. Liczba masowa jądra nie ulega zmianie, a ładunek wzrasta o dwa. Z jądra wylatują dwa elektrony i dwa antyneutrino. Jeżeli układ dwu antyneutrino jest identyczny z układem neutrino-antyneutrino, to może anihilować, co prowadzi do podwójnego, bezneutrinowego rozpadu β , przy którym z jądra wylatują tylko dwa elektrony. Odkrycie tego efektu byłoby dowodem, że neutrino są cząstkami Majorany⁵. Poszukiwania trwają, ale są trudne, bo prawdopodobieństwo takiego rozpadu jest bardzo małe. To prawdopodobieństwo szacuje się teoretycznie, ale rachunki wymagają robienia nie całkiem pewnych założeń z zakresu fizyki jądrowej, więc wyniki mogą być zawodne. Na razie nie ma przekonującego dowodu doświadczalnego, że takie rozpady istnieją. Fundamentalny problem, czy neutrino są cząstkami Majorany, pozostaje otwarty.

Z punktu widzenia teorii ważne jest, że jeżeli neutrino jest cząstką Majorany i jeżeli funkcja $\psi(x)$ opisuje jakiś stan neutrino, to sprzężona do niej ładunkowo funkcja $\psi^C(x)$ też opisuje stan neutrino, na ogół zresztą inny niż stan opisany funkcją $\psi(x)$. Przestrzeń funkcji falowych neutrino jest więc niezmiennicza nie tylko względem właściwych transformacji Lorentza, ale także względem sprzężenia ładunkowego.

Ćwiczenie Pokazać, że dla dowolnej funkcji falowej $\psi(x)$ funkcje

$$\frac{1}{\sqrt{2}}(\psi(x) + \psi^C(x)), \quad \frac{1}{i\sqrt{2}}(\psi(x) - \psi^C(x))$$

nie zmieniają się przy sprzężeniu ładunkowym i że z takich funkcji można zbudować bazę w przestrzeni stanów cząstki Majorany. Wyjaśnić dlaczego nie można tak rozumować dla cząstek, które nie są cząstkami Majorany.

• • • • •

Ponieważ sprzężenie ładunkowe dodaje lub usuwa kropkę nad wskaźnikiem spinora, dla cząstek Majorany można napisać równania, w których, jak w równaniu Diraca, masa cząstki jest różna od zera. Interpretacja fizyczna tych równań wymaga jednak kwantowej teorii pola, nie będziemy więc ich tu omawiać. Więcej informacji o teorii Majorany można znaleźć w artykule [18].

⁵Istnieje i taka możliwość, że podwójny, bezneutrinowy rozpad β będzie dominowany przez jakiś efekt nie wynikający z obecnej teorii. Poszukiwanie "nowej fizyki" jest jedną z motywacji badania procesów, które według znanej fizyki są bardzo mało prawdopodobne.

Bibliografia

- [1] K. Zalewski, *Wykłady z nierelatywistycznej mechaniki kwantowej* PWN Warszawa (1997).
- [2] B.F. Schutz, *Wstęp do ogólnej teorii względności* PWN Warszawa (2002).
- [3] A.K. Wróblewski, *Historia Fizyki* PWN Warszawa (2006).
- [4] A. Staruszkiewicz, *Algebra i Geometria* Kraków (1993).
- [5] V.F. Weisskopf, *The visual appearance of rapidly moving objects* *Physics Today* **13**(1960)24.
- [6] Ch. Misner, K.S. Thorn i J.A. Wheeler, *Gravitation* Freeman, San Francisco, (1973).
- [7] F. Halzen i A.D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics* John Wiley and Sons, New York (1984).
- [8] S. Weinberg, *Teoria pól kwantowych* PWN Warszawa (1999) tom I.
- [9] J.D. Jackson, *Elektrodynamika klasyczna* PWN Warszawa (1987).
- [10] J. Łopuszański, *Rachunek spinorów* PWN Warszawa 1985.
- [11] J.J. Sakurai, *Invariance Principles and Elementary Particles* Princeton University Press, Princeton, New Jersey (1964).
- [12] J.R. Taylor, *Scattering Theory* John Wiley and Sons, New York, London, Sydney, Toronto (1972).
- [13] J.J. Sakurai, *Advanced Quantum Mechanics* Adison Wesley (1967).
- [14] P.Krekora, Q. Su and R. Grobe, *Klein paradox in spatial and temporal resolution* *Phys.Rev.Letters* **92**(2004)040406.
- [15] J. Beringer et al. (Particle Data Group), *Review of Particle Physics* *Phys. Rev* **D86**(2012)010001.
- [16] V.B. Berestecki, E.M. Lifszyc i L.P. Pitajewski, *Elektrodynamika kwantowa* PWN Warszawa (2011).
- [17] R. Buder, *Radar Wynalazek, który zmienił Świat* Prószyński i ska, Warszawa.
- [18] F. Wilczek, *Majorana returns* *Nature Physics*, **5**(2009)614.